Title

Psychometric Performance and Responsiveness of the Functional Outcomes of Sleep Questionnaire and Sleep Apnea Quality of Life Index in a Randomized Trial: The HomePAP Study

Journal

Sleep, 37(12)

ISSN

0161-8105

Authors

Billings, Martha E
Rosen, Carol L
Auckley, Dennis
et al.

Publication Date

2014-12-01

DOI

10.5665/sleep.4262

Peer reviewed

Psychometric Performance and Responsiveness of the Functional Outcomes of Sleep Questionnaire and Sleep Apnea Quality of Life Index in a Randomized Trial: The HomePAP Study

Martha E. Billings, MD, MSc[1]; Carol L. Rosen, MD[2,3]; Dennis Auckley, MD[2,4]; Ruth Benca, MD, PhD[5]; Nancy Foldvary-Schaefer, DO, MS[2,6]; Conrad Iber, MD[7]; Phyllis C. Zee, MD, PhD[8]; Susan Redline, MD, MPH[9]; Vishesh K. Kapur, MD, MPH[1]

[1]UW Medicine Sleep Center, Division of Pulmonary Critical Care, University of Washington, Seattle, WA; [2]Case Western Reserve University, School of Medicine, Cleveland, OH; [3]Rainbow Babies and Children's Hospital, University Hospitals, Case Medical Center, Cleveland, OH; [4]Center for Sleep Medicine, Division of Pulmonary, Critical Care and Sleep Medicine, MetroHealth Medical Center, Cleveland, OH; [5]Department of Psychiatry, University of Wisconsin, Madison, WI; [6]Cleveland Clinic, Cleveland, OH; [7]Hennepin County Medical Center, University of Minnesota, Minneapolis, MN; [8]Northwestern University, Feinberg School of Medicine, Chicago, IL; [9]Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

Study Objectives: Measures of health-related quality of life (HRQL) specific for sleep disorders have had limited psychometric evaluation in the context of randomized controlled trials (RCTs). We investigated the psychometric properties of the Functional Outcomes of Sleep Questionnaire (FOSQ) and Sleep Apnea Quality of Life Index (SAQLI). We evaluated the FOSQ and SAQLI construct and criterion validity, determined a minimally important difference, and assessed for associations of responsiveness to baseline subject characteristics and continuous positive airway pressure (CPAP) adherence in a RCT population.

Design: Secondary analysis of data collected in a multisite RCT of home versus laboratory-based diagnosis and treatment of obstructive sleep apnea (HomePAP trial).

Participants: Individuals enrolled in the HomePAP trial (n = 335).

Interventions: N/A.

Measurement and Results: The FOSQ and SAQLI subscores demonstrated high reliability and criterion validity, correlating with Medical Outcomes Study 36-Item Short Form Survey domains. Correlations were weaker with the Epworth Sleepiness Scale (ESS). Both the FOSQ and SAQLI scores improved after 3 mo with CPAP therapy. Averaging 4 h or more of CPAP use was associated with an increase in the FOSQ beyond the minimally important difference. Baseline depressive symptoms and sleepiness predicted FOSQ and SAQLI responsiveness; demographic, objective obstructive sleep apnea (OSA) severity and sleep habits were not predictive in linear regression.

Conclusions: The Functional Outcomes of Sleep Questionnaire (FOSQ) and Sleep Apnea Quality of Life Index (SAQLI) are responsive to CPAP intervention, with the FOSQ being more sensitive to differences in CPAP adherence than the SAQLI. These instruments provide unique information about health outcomes beyond that provided by changes in physiological measures of OSA severity (apnea-hypopnea index).

Clinical Trial Information: Portable Monitoring for Diagnosis and Management of Sleep Apnea (HomePAP) URL: http://clinicaltrials.gov/show/NCT00642486. NIH clinical trials registry number: NCT00642486.

Keywords: health-related quality of life, psychometrics, sleep apnea

Citation: Billings ME, Rosen CL, Auckley D, Benca R, Foldvary-Schaefer N, Iber C, Zee PC, Redline S, Kapur VK. Psychometric performance and responsiveness of the Functional Outcomes of Sleep Questionnaire and Sleep Apnea Quality of Life Index in a randomized trial: the HomePAP study. *SLEEP* 2014;37(12):2017-2024.

INTRODUCTION

Disrupted sleep and daytime sleepiness, common symptoms of obstructive sleep apnea (OSA) in adults, are linked to decreased sense of well-being.[1] Subjects with sleep apnea often, although not uniformly, report impaired quality of life, especially in areas of vitality.[2] The severity of OSA as determined physiologically by the apnea-hypopnea index (AHI) or hypoxemic burden has not typically correlated with generic measures of health-related quality of life (HRQL)[3] or sleep-specific HRQL instruments.[4–9] Thus, it is difficult to predict the effect of OSA on patient- reported outcomes based on physiological severity. Well-validated, sleep-specific HRQL instruments are important for evaluating patient reported outcomes. Several instruments have been developed to assess the consequences of poor sleep and daytime sleepiness caused by sleep disorders on role functioning,[10] including the Functional Outcomes of Sleep Questionnaire (FOSQ)[11] and the Calgary Sleep Apnea Quality of Life Index (SAQLI).[12] The SAQLI was designed specifically for OSA, whereas the FOSQ was constructed to measure the consequences of sleep disorders in general.

Neither the FOSQ nor the SAQLI have been extensively validated in randomized control trial settings. Initial FOSQ instrument development used a convenience sample of individuals recruited from a sleep clinic visit to demonstrate test-retest reproducibility and internal reliability.[11] The FOSQ subscales were subsequently shown to correlate with the Medical Outcomes Study 36-Item Short Form Survey (SF-36)[13] domains using a sample of 51 OSA subjects participating in a research project.[11] The SAQLI has been shown to have high internal reliability. When concurrent validity was assessed using the SF-36 in an initial validation study of 24 patients with OSA, it showed high correlation only with the SF-36 vitality domain.[12] The

Submitted for publication August, 2013
Submitted in final revised form June, 2014
Accepted for publication July, 2014

Address correspondence to: Martha E. Billings, MD, MSc, UW Medicine Sleep Center, Division of Pulmonary Critical Care Medicine, University of Washington, Harborview Medical Center, 325 Ninth Avenue, Box 359803, Seattle, WA 98104; Tel: (206) 744-4999; Fax: (206) 744-8584; E-mail: mebillin@uw.edu

SAQLI has discriminative ability for OSA: SAQLI scores were higher in those with severe sleep apnea compared to simple snorers.[12] The FOSQ has been widely used in sleep research studies,[6,14–17] but data on its response properties have not been uniformly reported.[4,18] The SAQLI is less widely used; to our knowledge, no further validation of the scale has been done.[4] Thus, we sought to assess the psychometric properties of the FOSQ and SAQLI in subjects enrolled in an OSA treatment trial and propose a minimally important difference (MID) score change for both instruments.

Continuous positive airway pressure (CPAP) has been shown to improve some measures of HRQL,[19] but does not consistently improve generic HRQL instrument scores.[1,4,18,20] The effect of CPAP treatment on FOSQ scores has been positive,[19,21] but the magnitude of benefit has been variable and at times with small clinical effect.[4] CPAP adherence may be associated with HRQL changes: greater use of CPAP would be expected to yield greater improvement in HRQL. Two prior studies have examined the dose of CPAP (hours/night) necessary to "normalize" the FOSQ to 18 or higher. A score of less than 18 is considered abnormal, reflecting a negative effect on sleepiness on quality of life. One study identified a threshold of 7 h,[22] but this was not seen in a second randomized study of CPAP.[23] In both studies, FOSQ score variations were large and responses were not well correlated with CPAP use. Prior studies did not adjust for differences in baseline factors, which may affect HRQL and responsiveness. Because HRQL could be adversely affected by comorbid insomnia, depression, or limited sleep opportunity, these confounding factors may limit the effect of CPAP therapy on HRQL. Further consideration of these factors may help elucidate the sources of variation in HRQL responses to CPAP therapy in patients with OSA.

The effect of sleep disorders and sleepiness on HRQL is likely mediated by cultural, social, economic, and individual factors. As defined by the World Health Organization, HRQL reflects the individual's perception of his or her life in the context of their culture and value system, in relation to individual goals, expectations, standards, and concerns.[24] HRQL is by nature a subjective rating incorporating physical well-being, psychological health, social relationships, and interactions with the environment that produce a sense of gratification with life. As such, we hypothesized that the FOSQ and SAQLI baseline scores and responsiveness to CPAP would be associated with sociodemographics (sex, race, education, and residential socioeconomic status) and sleep opportunity. We also hypothesized that greater CPAP adherence would result in larger instrument response, reflecting an index of "dose" of the intervention, after adjusting for individual factors.

METHODS

We used patient reported outcomes data obtained from a randomized trial, the HomePAP study. As described elsewhere,[25] the HomePAP study compared unattended home-based studies to laboratory-based testing for the diagnosis and treatment of sleep apnea in subjects with moderate to severe OSA and a minimum Epworth Sleepiness Scale (ESS)[26] score of 12. The primary outcome for the study was CPAP adherence at 3 mo. The study demonstrated noninferiority of the home-based arm compared to the laboratory arm; subjects in the home arm had better adherence overall. Subjects completed questionnaires assessing demographic information, sleep habits as well as the ESS, SF-36,[13] Center for Epidemiologic Studies Depression Scale (CES-D),[27] and FOSQ instruments at enrollment using pen and paper. The SAQLI was administered by the study coordinator at each site. Subjects were randomized to either home or laboratory-based pathways. Subjects verified to have an AHI of $\geq$ 15 qualified to continue in the trial and were started on CPAP after home or laboratory-based titration. At 3 mo, data on CPAP use were downloaded and subjects completed the same sleep habits questionnaires and instruments.

Instruments

The FOSQ is a 30-item self-administered questionnaire assessing the effect of excessive sleepiness on activities of daily life.[11] It includes five subscales identified through factor analysis: activity level, vigilance, intimacy, general productivity, and social outcome.[1] Subjects are asked if sleepiness interferes with performing a given task; responses range from 1 (extreme difficulty) to 4 (no difficulty). Responses are averaged (excluding missing responses) to create a subscale score of 1 to 4, and then subscale scores are summed (5-20 for the total score), with greater scores indicating less effect of sleepiness on daily life.[28] A "normal" score has been proposed to be higher than 18.[17]

The Calgary Sleep Apnea Quality of Life Index (SAQLI) is an interviewer-administered questionnaire evaluating the effect of sleep apnea on daily activities, emotional functioning, social interactions, and symptoms over the past 4 w. If subjects are receiving treatment for OSA, an additional subscale includes treatment-related symptoms. Subjects' responses range from 1, which indicates a large effect (e.g. "all-of-the-time"), to 7, which indicates no effect (e.g. "none-of-the-time") for 56 items. The four subscales results are averaged (1-7) and the total score is an average of the subscales (1-7). A greater SAQLI score indicates less of an effect of OSA on quality of life.

Analysis

To assess the internal reliability of the FOSQ and SAQLI in the HomePAP population, we calculated the Cronbach α coefficient for each of the five subscales of the FOSQ[10] and for the four SAQLI subscales as well as for the overall FOSQ and SAQLI total scores.[12] Data from the baseline examination were used for internal reliability assessment, and included all subjects enrolled (regardless of their subsequent eligibility for continuation in the trial). We also conducted these analyses on the sample eligible for CPAP with data for 3-mo follow-up.

To assess criterion validity, we assessed the correlation between baseline FOSQ and SAQLI subscales and SF-36 domains using the pairwise Pearson correlations because the scales had a relatively normal distribution. Construct validity was assessed by evaluating the correlation of the FOSQ and SAQLI scores with the ESS,[26] an extensively used measure of sleep propensity. We also evaluated the correlation of the scales with each other. We assessed for an association of the FOSQ and SAQLI summary scores with AHI severity. AHI categories—defined as 0-14.9, II = 15-23.9, III = 24-49.9 and IV = > 50—were evaluated using the Cochran-Armitage test for trend.

To assess the responsiveness of the FOSQ and SAQLI to CPAP therapy, we calculated the change in the FOSQ and

SAQLI subscores and summary scores from baseline to 3-mo follow-up, as well as computed the effect-size, the standardized response mean, and performed paired *t*-tests. We evaluated for differences in responsiveness of the FOSQ and SAQLI summary scores by CPAP adherence (> 4 h versus less), sex, race (white versus nonwhite), AHI quartile, education (more than high school versus high school or less), ZIP code SES (lowest quartile versus others), self-reported short sleepers (< 6 h versus others) and self-reported long sleep latency (> 30 min versus other) and depressive symptoms (CES-D > 16 versus less)[29] using two sample *t*-test and one-way analysis of variance. Residential SES was derived from multiple census SES measures of the subject's ZIP code as described previously.[30] We developed predictive models of FOSQ and SAQLI responsiveness by performing stepwise linear regression. We considered the following as possible explanatory variables: CPAP adherence (> 4 h), demographics, baseline sleepiness (ESS), body mass index (BMI), AHI quartile, depressive symptoms, sleep latency and duration, randomization assignment, and FOSQ and SAQLI total scores at baseline for inclusion in the model.

Finally, we used distribution-based methods to determine the MID in FOSQ and SAQLI scores because the study did not include a suitable anchor question. The standard error of measurement (SEM) reflects the error in the observed score compared to true score and is sample-independent. We used one SEM to represent a meaningful intraindividual HRQL instrument change.[31] We calculated the SEM using baseline values for the standard deviation (SD) and reliability among those who followed up at 3 mo with the formula: $SEM = SD \times \sqrt{(1 - \alpha)}$.

RESULTS

Sample

Baseline enrollment FOSQ, SAQLI and SF-36 questionnaires data were available for 335 subjects. At enrollment, the subjects were a mean age 46.6 y (SD 12.3), 61% male, with BMI mean 37 (SD 8.6), 41% had a college degree, and 66% were white; 21% were black (Table 1). Of the 335 subjects who completed their screening diagnostic study, 191 qualified for continuation in the study with an AHI ≥ 15 on their diagnostic polysomnography (PSG). One hundred thirty-five subjects (71%) returned for follow-up at 3 mo, completed subsequent FOSQ and SAQLI questionnaires and had CPAP adherence data. Compared to the characteristics from the original screening sample, subjects available at 3-mo follow-up were older, more obese, and had shorter sleep time, (P < 0.05 for differences) but did not differ statistically by baseline ESS, total FOSQ, or SAQLI scores. The SF-36 vitality scores were the lowest at baseline and improved the most substantially at 3 mo by a mean of 13.4 (95% confidence interval [CI] 11.6, 15.2); all domains improved over 3 mo (Table 2). Questionnaire completion rates were similar for all instruments with the exception of the FOSQ sexual intimacy items (missing n = 57). The FOSQ and SAQLI subscores were all significantly higher at 3 mo (Table 3).

Internal Reliability

When including all enrollees, the Cronbach α coefficient was 0.94 for the 30-item FOSQ total score. By subscales, the

Table 1—HomePAP study sample baseline characteristics.

	All enrollees (n = 335)	Only among those with follow-up (n = 135)
Age, mean (SD) [a]	46.6 (12.3)	49.5 (12.6)
Male, % (n)	60.6 (203)	64 (87)
BMI, mean (SD) [a]	37.2 (8.6)	38.6 (8.8)
AHI, mean (SD) [a]	28.1 (28)	45.6 (26.3)
ESS, mean (SD)	14.1 (3.7)	14.3 (3.7)
CES-D, mean (SD)	11.13 (5.69)	10.6 (6.1)
Sleep duration < 6 h, % (n) [a]	38.9 (145)	50.4 (68)
Sleep latency > 30 min, % (n) [a]	30.3 (113)	20.7 (28)
Race, % (n)		
White	66 (221)	70 (95)
Black	21 (71)	15 (21)
Hispanic	8 (27)	10 (13)
Other	5 (16)	4 (6)
Education, % (n)		
HS or less	21.5 (72)	18 (24)
> HS but < college	37 (124)	41 (55)
College or more	41.5 (139)	42 (56)

[a] Baseline difference in all enrollees vs. those with follow-up significant (P < 0.05). AHI, apnea-hypopnea index; BMI, body mass index; CES-D, Center for Epidemiologic Studies Depression Scale; ESS, Epworth Sleepiness Scale; HS, high school; SD, standard deviation.

Table 2—Baseline and 3-mo Medical Outcomes Study 36-Item Short Form Survey subscales scores in HomePAP trial participants, reported as mean (standard deviation).

SF-36 [a]	Baseline: all enrollees (n = 335)	Baseline: only those with follow-up (n = 135)	3-mo follow-up (n = 135)
Vitality	39.10 (9.84)	39.37 (10.36)	53.08 (9.84)
Social functioning	42.13 (10.94)	42.97 (11.50)	50.01 (9.37)
Role physical	41.84 (10.77)	41.21 (11.08)	49.12 (9.56)
Physical functioning	44.44 (10.45)	44.08 (10.31)	47.83 (10.22)
Role emotional	44.16 (12.10)	44.16 (12.88)	49.73 (9.42)
General health perception	43.00 (10.04)	42.46 (10.32)	46.81 (9.71)
Bodily pain	47.33 (10.58)	46.62 (10.43)	50.06 (10.37)
Mental health [b]	46.30 (10.40)	47.99 (10.53)	53.02 (9.19)

[a] Each Short Form-36 domain improvement significant (P < 0.001). [b] Significant difference between all enrollees and those with follow-up (P < 0.05).

reliabilities were: activity level 0.90 (nine items), vigilance 0.86 (seven items), intimacy 0.85 (four items), general productivity 0.87 (eight items), and social outcome 0.82 (two items). The SAQLI total score (56 items) had a reliability of 0.92. The Cronbach α for the SAQLI subscales were: daily functioning 0.89 (11 items), social interactions 0.88 (13 items), emotional interaction 0.91 (11 items) and symptoms 0.70 (21 items). The reliability was slightly higher when baseline responses were restricted to those with 3-mo follow-up results (Table 3 reliability).

Table 3—Functional Outcomes of Sleep Questionnaire and Sleep Apnea Quality of Life Index scores, mean (standard deviation) at baseline and 3 mo follow-up, reliability and calculated standard error of measurement or minimally important difference among subjects with 3 month follow-up data (n = 135).

	Baseline Mean (SD)	3 mo Mean (SD)	Reliability (Cronbach α)	SEM/MID
FOSQ scale				
Vigilance (7 items) [a]	2.76 (0.63)	3.60 (0.52)	0.86	0.416
Social outcome (2 items)	3.28 (0.76)	3.79 (0.45)	0.82	0.322
General productivity (8 items)	3.15 (0.63)	3.72 (0.42)	0.87	0.227
Activity level (9 items) [a]	2.63 (0.74)	3.45 (0.55)	0.90	0.234
Sexual intimacy (4 items) [a]	2.90 (0.98)	3.61 (0.67)	0.85	0.380
Total score (30 items) [a]	14.73 (3.06)	18.17 (2.28)	0.94	0.750
SAQLI scale				
Daily functioning (11 items) [a]	3.58 (1.18)	5.19(1.00)	0.91	0.354
Social interactions (13 items) [a]	4.65 (1.34)	6.08 (1.02)	0.89	0.444
Emotional functioning (11 items) [a]	2.30 (1.27)	5.18 (1.01)	0.93	0.336
Symptoms (21 items y/n) [a]	2.31 (1.14)	3.20 (1.67)	0.74	0.581
Total score (56 items) [a]	3.68 (1.03)	4.53 (1.03)	0.92	0.291

[a] All statistically significant improvement at 3 mo, (P < 0.0001). FOSQ, Functional Outcomes of Sleep Questionnaire; MID, minimally important difference; SD, standard deviation; SEM, standard error of measurement; SAQLI, Sleep Apnea Quality of Life Index.

Criterion Validity

Baseline FOSQ and SAQLI subscores were significantly correlated with all SF-36 domains, (P < 0.001; Table 4). The FOSQ activity level was moderately correlated with all SF-36 domains except bodily pain. The SAQLI emotional functioning subscale and SF-36 mental health domain were strongly correlated (r = 0.65). The SF-36 social functioning and vitality domains were moderately correlated with all SAQLI subscales (r = 0.40-0.58). The FOSQ and SAQLI subscores and summary scores were also significantly correlated with each other, especially the SAQL daily functioning and FOSQ activity level subscores (r = 0.65) (Table 5).

Construct Validity

The ESS was inversely moderately correlated with the FOSQ

Table 4—Pearson correlation coefficients for baseline Medical Outcomes Study 36-Item Short Form Survey domains with baseline FOSQ and SAQLI subscales in the HomePAP population (n = 333); all P < 0.001.

	FOSQ				
SF-36	**General productivity**	**Social outcomes**	**Activity level**	**Vigilance**	**Sexual intimacy**
Vitality	0.479	0.321	0.645	0.266	0.362
Social functioning	0.547	0.495	0.612	0.349	0.348
Role physical	0.507	0.420	0.583	0.274	0.374
Physical functioning	0.380	0.312	0.439	0.227	0.221
Role emotional	0.429	0.300	0.440	0.192	0.333
General health	0.383	0.332	0.524	0.221	0.290
Bodily pain	0.300	0.283	0.330	0.188	0.290
Mental health	0.370	0.215	0.410	0.138	0.272
	SAQLI				
SF-36	**Daily functioning**	**Social interactions**	**Emotional functioning**	**Symptoms**	
Vitality	0.568	0.405	0.462	0.409	
Social functioning	0.536	0.554	0.578	0.403	
Role physical	0.529	0.425	0.391	0.369	
Physical functioning	0.384	0.364	0.299	0.331	
Role emotional	0.462	0.507	0.575	0.310	
General health	0.431	0.440	0.445	0.248	
Bodily pain	0.269	0.276	0.282	0.166	
Mental health	0.406	0.466	0.650	0.305	

FOSQ, Functional Outcomes of Sleep Questionnaire; SAQLI, Sleep Apnea Quality of Life Index; SF-36, Medical Outcomes Study 36-Item Short Form Survey.

Table 5—Correlation of the FOSQ and SAQLI Subscales and summary scores (n = 333); all P < 0.001.

	SAQLI				
FOSQ	**Daily functioning**	**Social interactions**	**Emotional functioning**	**Symptoms**	**Total score**
Vigilance	0.377	0.336	0.220	0.346	0.387
Social outcome	0.403	0.375	0.301	0.308	0.420
General productivity	0.613	0.512	0.460	0.441	0.615
Activity level	0.651	0.520	0.511	0.479	0.657
Total score	0.588	0.519	0.441	0.445	0.605

FOSQ, Functional Outcomes of Sleep Questionnaire; SAQLI, Sleep Apnea Quality of Life Index.

Table 6—Baseline Epworth Sleepiness Scale Pearson correlation coefficients in the HomePAP population (n = 334), all P < 0.001; with Functional Outcomes of Sleep Questionnaire subscales and total score and Sleep Apnea Quality of Life Index subscales and total score.

FOSQ subscale	**ESS**
Vigilance	-0.531
Social outcome	-0.300
General productivity	-0.377
Activity level	-0.310
Sexual intimacy	-0.129
Total score	-0.410
SAQLI subscale	**ESS**
Daily functioning	-0.269
Social interactions	-0.197
Emotional functioning	-0.081
Symptoms	-0.251
Total score	-0.243

ESS, Epworth Sleepiness Scale; FOSQ, Functional Outcomes of Sleep Questionnaire; SAQLI, Sleep Apnea Quality of Life Index.

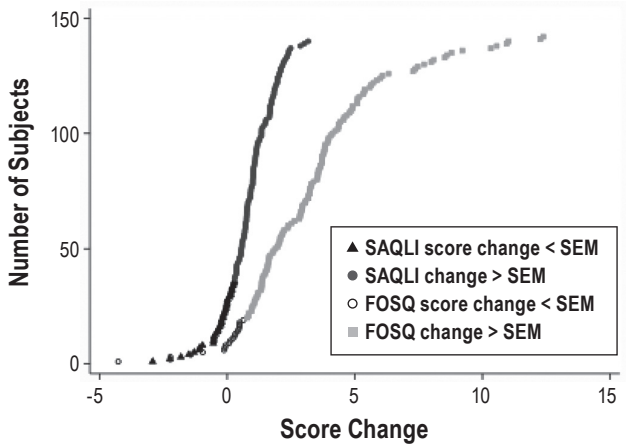

Figure 1—Plot of the cumulative distribution of Functional Outcomes of Sleep Questionnaire (FOSQ) and Sleep Apnea Quality of Life Index (SAQLI) change scores. The symbols mark the net difference in total score from baseline to 3 mo. SEM, standard error of measurement.

vigilance score (r = -0.56, P < 0.001; Table 6) as well as the FOSQ total score (r = -0.42; P < 0.001). The SAQLI subscores were all significantly correlated, albeit generally only weakly, with the ESS (Table 6).

Relation with AHI

The baseline FOSQ and SAQLI summary and subscale scores did not differ by AHI quartile, with the exception of the SAQLI emotional functioning subscore, which had a positive test for trend (P = 0.005) with higher scores in the higher AHI quartile. The AHI did not correlate with either the FOSQ or SAQLI summary or subscale scores, except weakly with the SAQLI emotional function (r = 0.134, P = 0.02).

Responsiveness

After 3 mo of CPAP therapy, individual FOSQ total scores increased by a mean of 3.3 points (95% CI: 2.86, 3.81), P < 0.001. The effect size was large at 1.15; the calculated SEM was 0.75 (Table 3). The FOSQ vigilance and activity level subscores had a greater responsiveness than the other subscores (Table 7). After 3 mo of CPAP, the SAQLI total score increased by a mean 0.82, (95% CI 0.65, 0.99), P < 0.001. The effect size was also large at 1.0; the SAQLI total score SEM was 0.29. The daily functioning and social interaction subscores had greater responsiveness than the others (Table 7). Using the 1 SEM cutoff, 86.6% of subjects were above the MID threshold for improvement in the FOSQ total score and 74.3% the SAQLI total score (Figure 1).

The FOSQ total score change was significantly higher in those CPAP adherent (> 4 h/day) but the SAQLI total score change was not. Those with a mean of 4 h or more of CPAP use per night (n = 63) increased their FOSQ total score by a mean of 3.9 (SD 2.8) compared to those averaging less than 4 h (n = 72), who had a mean increase of 2.9 (SD 2.8), P = 0.039, (Table 7). In bivariate analysis, the FOSQ total score responsiveness was also significantly greater in women, whites (versus nonwhites) and those with baseline depressive symptoms. The SAQLI total score change was greater in those with baseline depressive symptoms but did not differ by demographics. There was no difference in responsiveness by AHI quartile or sleep habits.

In stepwise regression models, CPAP use, baseline depressive symptoms, and sleepiness were predictive of FOSQ responsiveness (Tables 8 and 9). A baseline CES-D > 16 predicted of an increase of 3.1 (SD 0.6) in FOSQ and of 0.54 (SD 0.24) in SAQLI total score. Demographic variables, baseline sleep habits, BMI, and AHI quartile were not predictive (P > 0.10)

Table 7—Responsiveness of Functional Outcomes of Sleep Questionnaire and Sleep Apnea Quality of Life Index subscores to 3 mo of continuous positive airway pressure therapy in HomePAP population, n = 135.

FOSQ Subscore Δ Baseline – 3 mo	Mean (SD), IQR	≥ 4 h CPAP (n = 72)	< 4 h CPAP (n = 63)	P
Vigilance	0.81 (0.66), 0.39–1.14	0.94 (0.60)	0.74 (0.67)	0.064
Activity	0.80 (0.70), 0.23–1.31	0.94 (0.69)	0.67 (0.70)	0.027
Productivity	0.55 (0.57), 0.13–0.76	0.61 (0.55)	0.52 (0.59)	0.40
Social outcome	0.50 (0.75), 0.0–1.0	0.60 (0.82)	0.40 (0.67)	0.14
Intimacy (n = 104)	0.66 (0.85), 0.0–1.0	0.83 (0.84)	0.50 (0.84)	0.046
Total score	3.33 (2.85), 1.36–4.61	3.91 (2.8)	2.90 (2.8)	0.039
SAQLI Subscore Δ Baseline – 3 mo	**Mean (SD), IQR**	**≥ 4 h CPAP (n = 72)**	**< 4 h CPAP (n = 63)**	**P**
Daily functioning	1.59 (1.24), 0.8–2.4	1.80 (1.2)	1.40 (1.3)	0.070
Social interactions	1.39 (1.24), 0.6–2.2	1.42 (1.2)	1.44 (1.3)	0.89
Emotional functioning	0.92 (1.11), 0.2–1.4	0.93 (1.2)	0.99 (1.1)	0.78
Symptoms	0.87 (1.77), 0.4–2.2	0.73 (1.9)	1.0 (1.6)	0.35
Total score	0.82 (1.01), 0.3–1.5	0.91 (1.0)	0.79 (1.0)	0.49

CPAP, continuous positive airway pressure; FOSQ, Functional Outcomes of Sleep Questionnaire; SAQLI, Sleep Apnea Quality of Life Index; SD, standard deviation; IQR, Inter Quartile Range.

Table 8—Predictors of Functional Outcomes of Sleep Questionnaire total score change after 3 mo of continuous positive airway pressure use, stepwise multivariate regression, n = 134, $R^2 = 0.51$.

	β (SE)	95% CI	P
Depressive symptoms	3.09 (0.63)	1.84, 4.34	< 0.001
CPAP > 4 h	1.07 (0.43)	0.22, 1.92	0.02
ESS baseline	0.23 (0.06)	0.12, 0.35	< 0.001

CI, confidence interval; CPAP, continuous positive airway pressure; ESS, Epworth Sleepiness Scale; SE, standard error.

Table 9—Predictors of Sleep Apnea Quality of Life Index total score change after 3 months of continuous positive airway pressure use, stepwise multivariate regression, n = 133, $R^2 = 0.25$.

	β (SE)	95% CI	P
Depressive symptoms	0.54 (0.24)	0.50, 1.04	0.031
ESS baseline	0.04 (0.02)	-0.01, 0.08	0.11

CI, confidence interval; ESS, Epworth Sleepiness Scale; SE, standard error.

of FOSQ or SAQLI responsiveness. If baseline HRQL was included in the models, depressive symptoms and ESS were no longer predictors.

DISCUSSION

These intervention data from a clinical trial provide further evidence that the FOSQ and SAQLI are responsive to CPAP treatment for 3 mo. In a detailed assessment of several psychometric properties, we found that both the FOSQ and SAQLI demonstrated excellent internal reliability and had appropriate correlation with the well-established SF-36. The instruments demonstrated strong criterion validity as evidenced by moderate correlations with SF-36 social role and physical role domains. The social domains in SF-36, FOSQ, and SAQLI appear to measure similar constructs of social well-being. Surprisingly, the FOSQ vigilance was only weakly associated with the SF-36 vitality. The SAQLI showed much stronger correlations with the emotional role and mental health domains of the SF-36 than the FOSQ.

The FOSQ and SAQLI each were correlated with the ESS, albeit weakly. Other studies, one of community-dwelling elderly and another of sleep clinic patients, have found more robust association with the FOSQ and ESS, but with a similar stronger association of the vigilance scale and ESS.[28,32] Our correlations were likely impacted by our study cohort inclusion criterion of ESS > 12, constraining the distribution. Daytime sleepiness is a key subjective measure of OSA severity. The FOSQ, which was specifically developed to measure sleepiness related function, correlated with the ESS more than the SAQLI, which was designed to measure a broader impact of OSA on HRQL. Of interest, the FOSQ and SAQLI instrument scores did not differ by AHI quartile. A lack of association of the FOSQ and AHI is consistent with prior studies, and underscores the frequent weak association between objective measures of physiological disturbances and the subjective patient reported outcomes.[8,9,33]

The FOSQ demonstrated a slightly greater responsiveness to CPAP therapy. The FOSQ total score, activity, and vigilance subscores effect sizes (responsiveness to CPAP) were large and more substantial than in prior studies.[4,14] Over 4 h of CPAP use predicted greater improvement in the FOSQ total score. A dose response of greater FOSQ improvement with greater CPAP use was seen, as previously demonstrated by Weaver et al.[22] The mean FOSQ total score improvement of 1 unit (SD 0.43) associated with CPAP use is above the SEM or proposed MID. As the 4 h per night metric is typically used to define minimal adherence for therapeutic benefit, a 1-unit difference in the FOSQ may be significant clinically. Although the SAQLI did show

responsiveness to CPAP therapy, the change was not associated with the amount of CPAP use. Some of the observed HRQL responsiveness may represent a regression to the mean and not the use of CPAP *per se*.

The interpretation of the FOSQ and SAQLI responsiveness is not well defined; it remains undetermined what magnitude of change is clinically significant. We propose that an intra-individual change in the FOSQ and SAQLI of 1 SEM reflects a minimally important difference of clinical significance. The MID reflects the smallest score difference in a patient-reported outcome that is perceived by the patient as a change (harmful or beneficial).[34] Most subjects met these criteria, many reaching twofold to threefold above the SEM. Further psychometric evaluation is needed to guide the application and interpretation of these instruments in clinical trials and clinical practice. As the MID depends not only on the instrument but also the sample and contextual characteristics, using samples from subjects enrolled research trials with similar sleep characteristics as well as more diverse, "real world" clinic patients will provide a possible MID range and error estimate. Additional studies using anchor-based methods[35] and other distribution approaches in different populations[34] will be useful to confirm our proposed MID based on the SEM.

The FOSQ and SAQLI instruments' responsiveness to CPAP therapy did not differ by baseline demographics, self-reported sleep habits, or AHI. Baseline depressive symptoms were strong predictors of HRQL responsiveness and strongly correlated with baseline HRQL. Thus, the independent effect of depression cannot be evaluated. Responsiveness of these HRQL instruments does not appear to be affected by personal factors such as BMI, AHI, race, education, or residential SES.

Our study population and analysis have several limitations. Subjects all had a minimal ESS of 12, limiting the ability to assess the correlation in individuals with a full range of sleepiness scores, and limiting the ability to evaluate whether non-sleepy OSA subjects have distinct HRQL responses to treatment. The relatively modest sample size at 3 mo follow-up (135 subjects with CPAP data) may lead to a sampling biases. Because all subjects received CPAP therapy (no control arm), we cannot conclude the response was solely an effect of CPAP. The change may also reflect primarily depressive symptom improvement on HRQL ratings; the CES-D was very responsive to CPAP therapy. There is also a potential of residual AHI and other unmeasured confounders contributing to HRQL score changes. Although we conducted multiple comparisons, which increase the risk for a type I error, the overall consistency of the reported associations mitigates a concern about inappropriate inferences. The potential generalizability of the findings across the AHI range is supported by use of data from the enrolled subjects who both did not have moderate OSA as well as those with more severe OSA. The reliability and validity results were similar when only eligible subjects were included in the analysis (data not shown/Table 3).

In summary, our results indicate that in patients with OSA and a minimal ESS score of 12, both the FOSQ and SAQLI demonstrate good internal reliability, criterion validity with moderate correlation with the SF-36, and responsiveness to intervention. These instruments provide information distinct from the AHI. Furthermore, responsiveness of instruments did not differ by demographic factors or sleep habits supporting their use across diverse populations. Although the FOSQ demonstrated a slightly greater responsiveness to CPAP than the SAQLI, this may be because of the selection of patients with an elevated ESS on enrollment. Future research on OSA subjects with and without sleepiness is needed to confirm these findings. Nonetheless, our psychometric analyses support the utility and validity of both instruments for assessing key patient reported outcomes in OSA.

ABBREVIATIONS

AHI, apnea-hypopnea index
CES-D, Center for Epidemiologic Studies Depression Scale
CPAP, continuous positive airway pressure
ESS, Epworth Sleepiness Scale
FOSQ, Functional Outcomes of Sleep Questionnaire
HRQL, Health Related Quality of Life
OSA, obstructive sleep apnea
MID, Minimally Important Difference
RCT, randomized controlled trial
SAQLI, Sleep Apnea Quality of Life Index
SEM, standard error of measurement
SF-36, Medical Outcomes Study 36-Item Short Form Survey

ACKNOWLEDGMENTS

The authors thank Beth Devine, PhD, PharmD, MBA and Tracy Yep, PharmD for their assistance with the psychometric analysis and minimal important difference calculations.

DISCLOSURE STATEMENT

This was not an industry supported study. Support for this study was provided by the American Sleep Medicine Foundation 38-PM-07 Grant: Portable Monitoring for the Diagnosis and Management of OSA. Dr. Rosen has consulted for for Jazz Pharmaceutical. Dr. Auckley research support from Teva and ResMed. Dr. Benca has consulted for Sanofi-Aventis and Merck. Dr. Iber has provided uncompensated consultation services to Apnex Medical. Dr. Foldvary-Schaefer has received research support from CleveMed and UCB, Pharma and has participated in speaking engagements for Jazz Pharmaceuticals. Dr. Zee has consulted for Sanofi-Aventis, Merck, Philips-Respironics and Purdue. Dr. Redline has received research support from Philips-Respironics and ResMed Foundation. The other authors have indicated no financial conflicts of interest.

REFERENCES

1. Reimer MA, Flemons WW. Quality of life in sleep disorders. Sleep Med Rev 2003;7:335–49.
2. Moyer CA, Sonnad SS, Garetz SL, Helman JI, Chervin RD. Quality of life in obstructive sleep apnea: a systematic review of the literature. Sleep Med 2001;2:477–91.
3. Macey PM, Woo MA, Kumar R, Cross RL, Harper RM. Relationship between obstructive sleep apnea severity and sleep, depression and anxiety symptoms in newly-diagnosed patients. PLoS One 2010;5:e10211.
4. Jing J, Huang T, Cui W, Shen H. Effect on quality of life of continuous positive airway pressure in patients with obstructive sleep apnea syndrome: a meta-analysis. Lung 2008;186:131–44.
5. Montserrat JM, Ferrer M, Hernandez L, et al. Effectiveness of CPAP treatment in daytime function in sleep apnea syndrome: a randomized controlled study with an optimized placebo. Am J Respir Crit Care Med 2001;164:608–13.
6. Monasterio C, Vidal S, Duran J, et al. Effectiveness of continuous positive airway pressure in mild sleep apnea-hypopnea syndrome. Am J Respir Crit Care Med 2001;164:939–43.

7. Barbe F, Mayoralas LR, Duran J, et al. Treatment with continuous positive airway pressure is not effective in patients with sleep apnea but no daytime sleepiness. a randomized, controlled trial. Ann Intern Med 2001;134:1015–23.
8. Kezirian EJ, Harrison SL, Ancoli-Israel S, et al. Behavioral correlates of sleep-disordered breathing in older women. Sleep 2007;30:1181–8.
9. Kezirian EJ, Harrison SL, Ancoli-Israel S, et al. Behavioral correlates of sleep-disordered breathing in older men. Sleep 2009;32:253–61.
10. Devine EB, Hakim Z, Green J. A systematic review of patient-reported outcome instruments measuring sleep dysfunction in adults. PharmacoEconomics 2005;23:889–912.
11. Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. Sleep 1997;20:835–43.
12. Flemons WW, Reimer MA. Development of a disease-specific health-related quality of life questionnaire for sleep apnea. Am J Respir Crit Care Med 1998;158:494–503.
13. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–83.
14. Barnes M, Houston D, Worsnop CJ, et al. A randomized controlled trial of continuous positive airway pressure in mild obstructive sleep apnea. Am J Respir Crit Care Med 2002;165:773–80.
15. Barnes M, McEvoy RD, Banks S, et al. Efficacy of positive airway pressure and oral appliance in mild to moderate obstructive sleep apnea. Am J Respir Crit Care Med 2004;170:656–64.
16. Marshall NS, Neill AM, Campbell AJ, Sheppard DS. Randomised controlled crossover trial of humidified continuous positive airway pressure in mild obstructive sleep apnoea. Thorax 2005;60:427–32.
17. Drummond F, Doelken P, Ahmed QA, et al. Empiric auto-titrating CPAP in people with suspected obstructive sleep apnea. J Clin Sleep Med 2010;6:140–5.
18. Weaver TE. Outcome measurement in sleep medicine practice and research. Part 1: assessment of symptoms, subjective and objective daytime sleepiness, health-related quality of life and functional status. Sleep Med Rev 2001;5:103–28.
19. Gay P, Weaver T, Loube D, Iber C. Evaluation of positive airway pressure treatment for sleep related breathing disorders in adults. Sleep 2006;29:381–401.
20. McDaid C, Griffin S, Weatherly H, et al. Continuous positive airway pressure devices for the treatment of obstructive sleep apnoea-hypopnoea syndrome: a systematic review and economic analysis. Health Technol Assess 2009;13:iii-iv, xi-xiv, 1–119, 43–274.
21. Weaver TE, Mancini C, Maislin G, et al. Continuous positive airway pressure treatment of sleepy patients with milder obstructive sleep apnea: results of the CPAP Apnea Trial North American Program (CATNAP) randomized clinical trial. Am J Respir Crit Care Med 2012;186:677–83.
22. Weaver TE, Maislin G, Dinges DF, et al. Relationship between hours of CPAP use and achieving normal levels of sleepiness and daily functioning. Sleep 2007;30:711–9.
23. Antic NA, Catcheside P, Buchan C, et al. The effect of CPAP in normalizing daytime sleepiness, quality of life, and neurocognitive function in patients with moderate to severe OSA. Sleep 2011;34:111–9.
24. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. Soc Sci Med 1998;46:1569–85.
25. Rosen CL, Auckley D, Benca R, et al. A multisite randomized trial of portable sleep studies and positive airway pressure autotitration versus laboratory-based polysomnography for the diagnosis and treatment of obstructive sleep apnea: the HomePAP study. Sleep 2012;35:757–67.
26. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991;14:540–5.
27. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1:385–401.
28. Gooneratne NS, Weaver TE, Cater JR, et al. Functional outcomes of excessive daytime sleepiness in older adults. J Am Geriatr Soc 2003;51:642–9.
29. Boyd JH, Weissman MM, Thompson WD, Myers JK. Screening for depression in a community sample. Understanding the discrepancies between depression symptom and diagnostic scales. Arch Gen Psychiatry 1982;39:1195–200.
30. Billings ME, Auckley D, Benca R, et al. Race and residential socioeconomics as predictors of CPAP adherence. Sleep 2011;34:1653–8.
31. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. J Clin Epidemiol 1999;52:861–73.
32. Chasens ER, Sereika SM, Houze MP, Strollo PJ. Subjective and objective appraisal of activity in adults with obstructive sleep apnea. J Aging Res 2011;2011:751819.
33. Khan A, Harrison SL, Kezirian EJ, et al. Obstructive sleep apnea during rapid eye movement sleep, daytime sleepiness, and quality of life in older men in Osteoporotic Fractures in Men (MrOS) Sleep Study. J Clin Sleep Med 2013;9:191–8.
34. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61:102–9.
35. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 2003;56:395–407.