

Validity of Korean Version of Functional Outcomes of Sleep Questionnaire in Patients with Simple Snoring and Obstructive Sleep Apnea

Noh Eul Han, MA, Deok-Yong Kim, MA, Sang-Ahm Lee, MD

Department of Neurology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

Background and Objective We developed a Korean version of the Functional Outcomes of Sleep Questionnaire (K-FOSQ) and investigated its reliability and validity in simple snorer or obstructive sleep apnea (OSA) patients.

Methods A total 432 participants (70% men, 84% OSA, mean age 50.0 ± 9.8 years) who were simple snorers or had OSA were included. We assessed the internal consistency, test-retest reliability, factor analysis, multitrait scaling analysis, and the concurrent validity of the K-FOSQ. Participants completed a battery of questionnaires including the Epworth Sleepiness Scale (ESS), Short Form-36 Health Survey (SF-36), Medical Outcomes Study-Sleep (MOS-Sleep) Scale, and Beck Depression Inventory (BDI).

Results Factor analysis identified five factors, in which only 24 items met the loading criteria. The five factors of K-FOSQ accounted for 73.0% of the variance. Cronbach's alpha coefficient for all domains exceeded the 0.70 standard for internal consistency. Test-retest reliability was acceptable ($r = 0.41-0.93$). Item-domain correlations ranged from 0.37 to 0.90. Only one item did not reach the threshold of 0.40. Floor effects were not observed, but ceiling effects were marked on all K-FOSQ subscales except one. All domains of K-FOSQ were significantly correlated with the corresponding scores of all tested instruments. The global K-FOSQ had a strong correlations ($r > 0.50$) with ESS and Sleep Problem Index-2 of MOS-Sleep, and had medium-sized correlations ($r = 0.40-0.50$) with BDI and SF-36 total scores. The K-FOSQ global and subscales did discriminate between participants with and without daytime sleepiness, but not between simple snorers and OSA patients.

Conclusions The K-FOSQ is a reliable and valid instrument for assessing functional outcome in participants with daytime sleepiness.

Sleep Med Res 2014;5(1):5-14

Key Words Functional Outcomes of Sleep Questionnaire, Obstructive sleep apnea, Simple snoring, Daytime sleepiness, Epworth Sleepiness Scale.

INTRODUCTION

Obstructive sleep apnea (OSA) is one of the most common sleep disorders. OSA is characterized by repeated episodes of airflow cessation or reduction that result in episodes of nocturnal hypoxemia and repetitive arousal leading to fragmented, nonrestorative sleep.^{1,2} Patients with OSA frequently complain of daytime sleepiness,³ which predisposes them for accidents, interpersonal problems, and reduced productivity, and may also lead to deterioration in psychosocial and cognitive function.⁴ Daytime sleepiness is usually measured by the Epworth Sleepiness Scale (ESS),⁵ which is a useful clinical tool for subjectively assessing one's propensity to fall asleep and for quantifying the severity of excessive sleepiness.

The Functional Outcomes of Sleep Questionnaire (FOSQ)⁶ is a self-administered, condition-specific questionnaire designed for use in patients with sleep disorders. It was developed in the USA to evaluate the impact of excessive sleepiness on activities of daily living. The FOSQ consists of 30 items within five domains including Activity Level, Vigilance, Intimacy and Sexual Relationships, General Productivity, and Social Outcome. These diverse aspects of the FOSQ could complement the evaluation of sleepiness provided by the ESS.

Received: October 21, 2014
Revised: November 21, 2014
Accepted: December 2, 2014

Correspondence

Sang-Ahm Lee, MD
Department of Neurology,
Asan Medical Center,
University of Ulsan College of Medicine,
88 Olympic-ro 43-gil, Songpa-gu,
Seoul 138-736, Korea
Tel +82-2-3010-3445
Fax +82-2-474-4691
E-mail salee@amc.seoul.kr

To be used in a different language and cultural setting, the questionnaire should be translated and adapted cross-culturally and then should be validated to be equivalent to the original instrument. The purposes of this study were 1) to examine the factor structure of the Korean translated version of FOSQ (K-FOSQ) and 2) to examine the reliability and validity of the K-FOSQ in patients with simple snoring and OSA.

METHODS

Subjects

Participants were adult patients who visited a sleep laboratory for an evaluation of suspected OSA in the Asan Medical Center from January 2011 to December 2012. Their primary language was Korean. Inclusion criteria of participation in the present study were as follows: over 18 years of age, a diagnosis of simple snoring or OSA after undergoing standard polysomnography (PSG), and having completed a battery of sleep- and health-related questionnaires. Patients were excluded if they had a diagnosis of any coexisting sleep disorder by history or PSG, if they had self-reported any medical or psychiatric illness, or they used sedative or hypnotic medications. As shown in Table 1, a total 432 participants (301 men and 131 women) were included in this study. The mean age of participants was 50.0 ± 9.8 years, 84% were patients with OSA, and the patients' mean body mass index was 26.0 ± 4.0 kg/m². ESS score was ≥ 11 in 164 patients (38.0%). Written informed consent was obtained from all patients.

Obstructive sleep apnea was diagnosed and evaluated using the standard PSG. An apnea was defined as a drop in the peak thermal sensor excursion of $\geq 90\%$ of the baseline value for at least 10 s.⁷ A hypopnea was defined as a nasal pressure signal excursion drop of $\geq 30\%$ of the baseline value for at least 10 s, accompanied by a $\geq 4\%$ reduction in O₂ saturation from the pre-event baseline. The apnea-hypopnea index (AHI) was defined as the sum of apneas and hypopneas per hour. Total sleep time was defined as the total number of minutes featuring exclusively N1, N2, N3, or REM sleep. Sleep stages were identified for each 30-second epoch by a well-trained registered polysomnographic technologist.

On the night of the PSG, patients completed a battery of sleep-related questionnaires. Basic demographic information, medical comorbidity, and medication information were obtained from these questionnaires and from the patient's electronic medical record.

Korean Version of Functional Outcomes of Sleep Questionnaire

The FOSQ is a self-administered questionnaire, and consists of 30 items focusing on the impact of sleep on 5 domains: Activity Level (9 items), Vigilance (7 items), Intimacy and Sexual Relationships (4 items), General Productivity (8 items), and Social Outcome (2 items).⁶ Each item is scored on a 4-point scale ranging from 1 to 4. An item that results in a score of 0 is coded as non-available or as a missing response which will not be included in the calculation. The range of scores is 1 to 4 for each of the 5 domains and 5 to 20 for the global (summed)

Table 1. Patient characteristics

	0 ≤ AHI < 10 (n = 113)	10 ≤ AHI < 30 (n = 148)	AHI ≥ 30 (n = 171)	p-value
Male, n (%)	44 (38.9)	107 (72.3)	150 (87.7)	< 0.001
Age, years	47.8 ± 11.0	51.8 ± 8.6	49.8 ± 9.6	0.004
Body mass index, kg/m ²	23.8 ± 3.4	25.9 ± 3.6	27.6 ± 4.0	< 0.001
Neck circumference, cm	35.6 ± 3.9	38.9 ± 3.4	40.9 ± 3.7	< 0.001
Apnea-hypopnea index, /h	4.1 ± 2.9	19.2 ± 5.4	52.1 ± 17.2	< 0.001
Respiratory distress index, /h	13.9 ± 7.3	28.7 ± 7.8	56.1 ± 15.4	< 0.001
Epworth Sleepiness Scale	8.8 ± 5.0	10.0 ± 5.0	10.1 ± 5.4	0.097
Short Form-36	69.1 ± 18.5	72.4 ± 19.4	72.0 ± 17.0	0.370
Sleep Problem Index-2	34.9 ± 18.0	35.2 ± 17.3	35.1 ± 16.9	0.990
Beck Depression Inventory	11.8 ± 7.2	11.3 ± 8.4	10.6 ± 6.9	0.458
K-FOSQ				
Vigilance	3.55 ± 0.56	3.53 ± 0.61	3.46 ± 0.62	0.401
General Productivity/Social Outcome	3.80 ± 0.36	3.85 ± 0.33	3.79 ± 0.39	0.314
Intimacy and Sexual Relationships	3.65 ± 0.62	3.60 ± 0.67	3.57 ± 0.71	0.665
Mental and Physical Activity Level	3.29 ± 0.54	3.37 ± 0.52	3.21 ± 0.62	0.054
Driving	3.54 ± 0.69	3.41 ± 0.73	3.39 ± 0.75	0.209
K-FOSQ global	17.8 ± 02.20	17.8 ± 02.26	17.4 ± 0.2.5	0.309

AHI: apnea-hypopnea index, K-FOSQ: Korean version of Functional Outcomes of Sleep Questionnaire.

score. Lower scores indicate greater daytime dysfunction.

We adapted the FOSQ into a Korean version as follows: we translated the FOSQ into Korean, conducted an assessment of item comprehension, performed a back-translation into English, and then developed a resulting version via consensus. Translation of the FOSQ into Korean was done by the corresponding author (Lee SA) and back-translation into English was done by a bilingual person. The consensus version was developed by the corresponding author (Lee SA).

Questionnaires Administered

Short Form-36 Health Survey

The Short Form-36 Health Survey (SF-36)⁸ assesses non-disease specific health-related quality of life. It contains 36 items covering 8 domains with four pertaining to physical functioning (physical functioning, role-physical, bodily pain, and general health) and four pertaining to mental health functioning (vitality, social functioning, role-emotional, and mental health). All raw scale score were linearly converted to a scale of 0–100. A higher score indicates better health-related functioning and quality of life. We used the Korean version of the SF-36.⁹

Epworth Sleepiness Scale

The ESS is the most widely used questionnaire to assess subjective daytime sleepiness.⁵ This is a self-administered scale with 8 items about how easily the respondent would fall asleep in different situations. The items are scored on a 0–3 scale, which are added to give an overall score of 0–24. A higher score indicates greater sleepiness during daily activities. An ESS score ≥ 11 is considered indicative of excessive daytime sleepiness. We used the Korean version of the ESS.¹⁰

Medical Outcomes Study-Sleep Scale

The Medical Outcomes Study-Sleep (MOS-Sleep) Scale^{11,12} is one of the most widely used scales for evaluating broad-spectrum sleep quality. It is a self-administered, non-disease specific scale for assessing information pertaining to both sleep quality and sleep quantity, consisting of 12 items. The MOS-Sleep measures the subjective experiences of sleep across 6 domains (sleep disturbance, 4 items; sleep adequacy, 2 items; sleep quantity, 1 item; daytime somnolence, 3 items; snoring, 1 item; and shortness of breath, 1 item). The Sleep Problems Index-2 (SPI-2) uses 9 items from 4 domains (sleep disturbance, 4 items; sleep adequacy, 2 items; daytime somnolence, 2 items; and shortness of breath, 1 item). Higher scores of SPI-2 indicate a more severe sleep problem. We used the Korean version of the MOS-Sleep Scale.¹³

Beck Depression Inventory

The Beck Depression Inventory (BDI) is one of the most commonly used scales to assess the severity of depressive symp-

toms. It is a 21-item self-administered scale.¹⁴ Each item contains 4 graded statements that reflect a 4-point scale ranging from 0 to 3. The scores range from 0 representing no depression to 63 representing severe depression. The Korean version of BDI has also been validated.¹⁵

Statistical Analysis

All statistical analyses were performed using Statistical Package for the Social Sciences (SPSS) 21 (SPSS Inc., Chicago, IL, USA). *p*-values < 0.05 were considered statistically significant.

Structural construct validity

Construct validity was assessed by exploratory factor analysis (principal axis factoring with promax rotation) to examine if the Korean version of the FOSQ captures the same constructs proposed in the original version. Exploratory factor analysis is a multivariate statistical method used to uncover the underlying structure of variable interrelations, and is a technique within factor analysis with an overarching goal to determine whether they can be represented by a smaller number of underlying factors. Criteria for extraction included: 1) loadings of at least 0.40 and at least 0.15 difference in cross-loadings, 2) use of the scree plot to identify the number of factors, and 3) eigenvalues greater than 1.0.¹⁶ When low factor loadings occurred, one item was deleted and the principal axis factoring was repeated until all items loaded according to the entry criteria.

To examine the appropriateness of the data for factor analysis, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and the Bartlett's Test of Sphericity were used. The KMO statistic varies from 0 to 1, and values higher than 0.7 are recommended.¹⁷

Reliability

Internal consistency of the K-FOSQ was assessed by the overall Cronbach's alpha coefficient and the Cronbach's alpha coefficient for each scale. Estimates of a magnitude higher than 0.7 were considered acceptable.¹⁸ Test-retest reliability was assessed using Pearson correlation coefficients. To examine test-retest reliability, an interval of two or three weeks between each assessment was chosen so as to minimize the subject's recall of the previous answer. Only patients without changes in their illness status were included. Therefore, the second K-FOSQ was obtained without intervening procedures (such as continuous positive airway pressure titration or sleep-related medication) when the subjects visited the outpatient clinic two or three weeks after PSG.

Multitrait scaling analysis

For the purpose of examining how well items of each domain represent a particular trait relative to other traits, item convergence and item discrimination were evaluated. Item convergence assesses the correlation between each item and its own

domain, and its criterion is met when the value is greater than 0.40.¹⁹ Item discrimination assesses the extent to which an item correlates more closely with the domain it represents than with other domains. Its criterion states that each item should have a higher correlation with its own domain than with any of the others.²⁰

Floor and ceiling effects

For floor and ceiling effects, we examined the proportion of scores with floor and ceiling effects using the total scale scores of the potential minimum and maximum, respectively. Ceiling and floor effects were below the acceptable cutoff of 15%.²¹

Concurrent validity

For concurrent validity, Spearman's rank correlation coefficients were calculated to assess the relationship between scores of K-FOSQ and other instruments administered in this study.

Discriminant validity

To assess this, we selected two types of parameters suggesting the severity and its consequence of OSA: the AHI, and scores of ESS. To examine the relations between or among them, Student's t test or a one-way analysis of variance test were used.

Table 2. Factor loadings in the rotated-factor matrix for the Korean version of Functional Outcomes of Sleep Questionnaire

Item	Factors				
	1	2	3	4	5
Factor 1: Vigilance (5 items)					
19. Difficulty enjoying concert	0.91	-0.06	-0.08	0.13	-0.01
18. Difficulty enjoying theater or lecture	0.88	-0.08	-0.03	0.14	0.00
17. Difficulty watching a movie	0.85	-0.05	0.05	-0.07	0.09
20. Difficulty watching television	0.78	-0.07	-0.02	-0.01	0.09
21. Difficulty participating in meetings of a group	0.50	0.15	0.04	0.13	0.11
Factor 2: General Productivity/Social Outcome (7 items)					
11. Difficulty maintaining a telephone conversation	-0.06	0.83	-0.06	-0.07	0.10
3. Difficulty finishing a meal	-0.27	0.81	-0.07	0.03	0.08
13. Difficulty visiting with family/friends in their home	0.38	0.73	0.04	-0.20	-0.10
12. Difficulty visiting with family/friends in your home	0.33	0.70	0.05	-0.21	-0.05
10. Difficulty performing employed or volunteer work	-0.18	0.66	-0.03	0.35	0.03
14. Difficulty doing things for family or friends	0.28	0.56	0.09	0.07	-0.16
9. Difficulty taking care of financial affairs and doing paperwork	0.05	0.42	0.00	0.26	0.09
Factor 3: Intimacy and Sexual Relationships (4 items)					
29. Ability to become sexually aroused affected	0.00	0.05	0.96	-0.04	-0.10
30. Ability to have an orgasm affected	0.10	-0.13	0.93	0.05	-0.09
28. Desire for intimacy or sex affected	-0.11	0.00	0.88	0.09	0.10
27. Intimate or sexual relationship affected	-0.07	0.00	0.81	0.02	0.18
Factor 4: Mental and Physical Activity Level (5 items)					
2. Difficulty remembering things	-0.01	0.02	0.00	0.79	0.02
4. Difficulty working on a hobby	-0.04	0.24	0.04	0.66	-0.11
1. Difficulty concentrating on things	0.20	-0.05	-0.12	0.62	0.14
26. Rating of general level of activity	0.05	-0.16	0.10	0.56	-0.11
23. Difficulty being as active as you want in the morning	0.14	0.04	0.11	0.50	0.01
Factor 5: Driving (3 items)					
6. Difficulty operating a motor vehicle for short distances	0.07	0.01	-0.03	-0.04	0.90
7. Difficulty operating a motor vehicle for long distances	0.08	-0.02	0.02	-0.07	0.86
8. Difficulty getting things done because too sleepy to drive	0.04	0.17	0.10	0.05	0.55
Eigenvalues	11.12	2.17	1.69	1.50	1.04
Cumulative variance (%)	46.31	9.02	7.02	6.26	4.35
Cronbach's alpha (total: 0.94)	0.92	0.87	0.94	0.80	0.86

RESULTS

Structural Constructive Validity

The factor analysis loadings yielded a five-factor solution that met the analytic criteria (Table 2). Principal axis factoring was computed on the 30-item FOSQ, and it indicated that 24 items (Appendix) met the loading criteria. Six items did not meet the loading criteria. Five items (items 15, 16, 22, 24, and 25) was deleted because they did not meet the loading criterion of > 0.40 , and one item (item 5) was deleted because it did not meet the criterion of a > 0.15 difference in cross-loading (0.47 on factor 1 vs. 0.49 on factor 4). After deleting the 6 items, five factors explaining a 73.0% of variance emerged. Loadings ranged from 0.50 to 0.91 for the first factor, 0.42 to 0.83 for the second factor, 0.81 to 0.96 for the third factor, 0.50 to 0.79 for the fourth factor, and 0.55 to 0.90 for the fifth factor. The screening test confirmed a five-factor solution. The KMO statistic was 0.921, supporting a finding that the data were suitable for factor analysis. The Bartlett's Test of Sphericity was significant ($\chi^2 = 4846.22$, $p < 0.001$), indicating that the correlation matrix was appropriate for the analysis. Table 2 shows factor loadings, eigenvalues, Cronbach alphas, and the cumulative variance for each factor.

The first factor consisted of five items from the original Vigilance domain, and explained 46.3% of the variance in the 26 items of the K-FOSQ. The second factor accounted for addi-

tional 9.0% of the variance. Seven items (items 3, 9, 10, and 11 from the original General Productivity domain, items 12 and 13 from the original Social Outcome domain, and item 14 from the original Activity Level domain) loaded on the second factor (named 'General Productivity/Social Outcome'). The third factor comprised all four items of the original Intimacy and Sexual Relationships domain, and explained additional 7.0% of the variance. The fourth factor counted for additional 6.3% of the variance. Five items (items 1, 2, and 4 from the original General Productivity domain and items 23 and 26 from the original Activity Level domain) loaded on the fourth factor (named 'Mental and Physical Activity Level'). The fifth factor explained an additional 4.4% of the total variance. Three items (items 6 and 7 from the original Vigilance domain and item 8 from the original General Productivity domain) loaded on the fifth factor (named 'Driving'). The deleted six items (items 5, 15, 16, 22, 24, and 25) from the FOSQ were all from the original Activity Level domain (Table 3).

Reliability

The analysis for reliability was performed on the 24 items and five domains yielded by the factor analysis. All domains showed good internal consistency in K-FOSQ (Appendix). Cronbach's alphas ranged from 0.80 (Mental and Physical Activity Level) to 0.94 (Driving) across five domains, and Cronbach's alpha was 0.94 for the total scale (Table 2).

In this study, test-retest reliability of K-FOSQ was assessed in 36 patients, and was acceptable. The correlation coefficient for the total was $r = 0.79$ ($p < 0.01$), and ranged from $r = 0.41$ to $r = 0.93$ in each domain (Factor 1, Vigilance, $r = 0.64$, $p < 0.01$; Factor 2, General Productivity/Social Outcome, $r = 0.80$, $p < 0.01$; Factor 3, Intimacy and Sexual Relationships, $r = 0.93$, $p < 0.01$; Factor 4, Mental and Physical Activity Level, $r = 0.81$, $p < 0.01$; Factor 5, Driving, $r = 0.41$, $p < 0.01$).

Multitrait Scaling Analysis

The items-to-domain correlations were calculated for 24 items comprising five domains. Item-domain correlations

Table 3. Items deleted from the original Functional Outcomes of Sleep Questionnaire

No.	Item
5	Difficulty doing work around the house
15	Relationship with family/friends affected
16	Difficulty exercising or participating in sport activity
22	Difficulty being as active as you want in evening
24	Difficulty being as active as you want in afternoon
25	Difficulty keeping pace with others your own age

Table 4. Item convergent and discriminant validity of the Korean version of FOSQ

Domains	Number of items	Item convergent validity		Item discriminant validity	%Floor	%Ceiling
		Range of correlations	Success rate (%)	Success rate (%)		
Vigilance	5	0.66–0.88	100	100	0.7	39.6
General Productivity/Social Outcome	7	0.37–0.74*	85.7	100	0	62.7
Intimacy and Sexual Relationships	4	0.83–0.90	100	100	1.4	57.6
Mental and Physical Activity Level	5	0.61–0.79	100	100	0.5	7.4
Driving	3	0.78–0.95	100	100	1.4	45.6
FOSQ global	24	-	-	-	0	4.9

*Only one item (item number 3) did not reach the threshold of 0.40.

%Floor: % with lowest possible score, %Ceiling: % with highest possible score, FOSQ: Functional Outcomes of Sleep Questionnaire.

ranged from 0.37 to 0.90 (Table 4). Only item 3 did not reach the threshold of 0.40. With regard to item discrimination, all items had a higher correlation with their own domains than they did with others.

Floor and Ceiling Effects

Table 4 showed the percentage of respondents giving the lowest (floor) and highest possible scores (ceiling). The percentages of participants with floor and ceiling effects for the global K-FOSQ were 0% and 4.9%, respectively. The K-FOSQ scores did not concentrate at the floor for any subscale. However, there were marked ceiling effects in all the K-FOSQ subscales except the Mental and Physical Activity Level subscale.

Concurrent Validity

Table 5 presents the correlation coefficients of K-FOSQ with the other instruments administered in this study. All domains of K-FOSQ significantly correlated with scores of all tested instruments. The total scores of K-FOSQ had particularly strong correlations ($r > 0.50$) with ESS and SPI-2 of MOS-Sleep, and had medium-sized correlations ($r = 0.40-0.50$) with the scores

from the BDI and the SF-36 total scores. Among the domains of K-FOSQ, the Mental and Physical Activity Level domain most significantly correlated with all tested instruments. The Vigilance and Driving domains were strong correlated with ESS ($r = -0.58, r = -0.53$) but were weakly correlated ($r < 0.40$) with other tested instruments (Table 5).

Discriminant Validity

Subjects were divided into 3 groups according to the severity of AHI (Table 1): the normal/mild group ($0 \leq \text{AHI} < 10$), the mild/moderate group ($10 \leq \text{AHI} < 30$), and the severe group ($\text{AHI} \geq 30$). Only scores of the Mental and Physical Activity domain tended to be different among the three subgroups ($p = 0.054$). The scores of the General Productivity/Social Outcome domain were significantly lower in the severe OSA subgroup (3.21 ± 0.62) had than the mild/moderate subgroup (3.37 ± 0.52) ($p = 0.042$). In addition, AHI tended to be weakly correlated with General Productivity/Social Outcome ($r = 0.094, p = 0.053$) and Vigilance ($r = -0.083, p = 0.085$). The FOSQ global and subscales did not discriminate between simple snorers and OSA patients.

Table 5. Spearman's rank correlation coefficients between K-FOSQ and other scales

	Vigilance	General Productivity/ Social Outcome	Intimacy and Sexual Relationships	Mental and Physical Activity Level	Driving	K-FOSQ global
SF-36 total	0.31**	0.39**	0.32**	0.50**	0.24**	0.42**
Physical Functioning	0.31**	0.27**	0.30**	0.32**	0.23**	0.35**
Role-physical	0.14**	0.22**	0.16**	0.25**	0.09	0.21**
Bodily Pain	0.20**	0.28**	0.12*	0.28**	0.13*	0.23**
General Health	0.15**	0.22**	0.23**	0.36**	0.15*	0.29**
Vitality	0.26**	0.32**	0.29**	0.46**	0.23**	0.39**
Social Functioning	0.20**	0.37**	0.28**	0.40**	0.21**	0.34**
Role-emotional	0.13*	0.16**	0.13*	0.19**	0.09	0.16**
Mental Health	0.23**	0.29**	0.21**	0.39**	0.19**	0.31**
Epworth Sleepiness Scale	-0.58**	-0.40**	-0.33**	-0.46**	-0.53**	-0.61**
Sleep Problem Index-2	-0.34**	-0.42**	-0.28**	-0.51**	-0.36**	-0.51**
Beck Depression Inventory	-0.34**	-0.38**	-0.33**	-0.44**	-0.26**	-0.44**

* $p < 0.05$. ** $p < 0.01$.

SF-36: Short Form-36 Health Survey, K-FOSQ: Korean version of the Functional Outcomes of Sleep Questionnaire.

Table 6. Differences in the K-FOSQ total between ESS score ≥ 11 and < 11

	ESS < 11	ESS ≥ 11	p-value
Vigilance	3.76 \pm 0.34	3.11 \pm 0.72	< 0.001
General Productivity/Social Outcome	3.91 \pm 0.20	3.66 \pm 0.50	< 0.001
Intimacy and Sexual Relationships	3.75 \pm 0.51	3.39 \pm 0.81	< 0.001
Activity Level	3.47 \pm 0.44	2.98 \pm 0.64	< 0.001
Driving	3.69 \pm 0.49	3.02 \pm 0.87	< 0.001
K-FOSQ global	18.6 \pm 1.38	16.2 \pm 2.77	< 0.001

K-FOSQ: Korean version of the Functional Outcomes of Sleep Questionnaire, ESS: Epworth Sleepiness Scale.

Subjects were also divided into 2 groups according to the severity of daytime sleepiness: the non-sleepy group ($ESS < 11$) and the sleepy group ($ESS \geq 11$). The scores of K-FOSQ total and five domains were significantly different between patients with and without daytime sleepiness (Table 6). As expected, patients with ESS scores ≥ 11 had significantly lower functional outcome than those with $ESS < 11$.

DISCUSSION

The application of the FOSQ in non-English-speaking countries requires linguistic adaptation together with a re-examination of its validity. Several language versions of the FOSQ including Norwegian,²² Spanish,²³ Swedish,²⁴ and Thai²⁵ have been recently evaluated in patients with sleep disorders, especially with OSA. So far, there is no published data of factor analysis for a non-English version of FOSQ. Factor analysis in this study constituted five factors which were consistent with the original FOSQ.⁶ However, only 24 questions were captured by the factor structure in the present study, and items consisting of individual factors were somewhat different from the original FOSQ. For example, three driving-related items from the Vigilance and General Productivity in the original FOSQ constituted the new Driving factor in this study. Additionally, the new Mental and Physical Activity Level is made up of three items from the General Productivity domain (difficulty concentrating on things, difficulty remembering things, and difficulty working on a hobby) and two items from the Activity Level domain (rating of general level of activity and difficulty being as active as you want in morning) in the original FOSQ. The Intimacy and Sexual Relationships domain was identical to the original one. All of deleted items in this study were from the original Activity Level domain. Cultural variations may account for this discrepancy. Cultural variability could seriously affect a questionnaire design and the expected outcomes.²⁶ The proportion of total variance in the set of 24 questions captured by the factor structure was 73.0%, which was higher than the 57.3% of the original FOSQ.

The internal consistency of K-FOSQ (Appendix) was found to be good to excellent in all domains (Cronbach's alpha 0.80 to 0.94) and for the global score (Cronbach's alpha 0.94). And in the test-retest the intraclass correlation coefficients varied between 0.41 and 0.93, with 0.79 for the global score. The lowest reproducibility was found in the Driving subscale. All domains and global scores except Driving showed acceptable test-retest reliability.

The K-FOSQ was also shown to be valid for measuring the concept of the hypothesized dimension. In this study, all items in domains except one item (difficulty finishing a meal) from the General Productivity/Social Outcome domain had higher item-scale correlations than 0.40 for the hypothesized dimension.

These findings represent good item correlations with their own domain.¹⁹ Item discrimination was also satisfied. The scaling success rates on discriminant validity were 100% for all domains. All items showed lower item correlations (less than 0.40) with other domains. This suggests that items of K-FOSQ were more strongly correlated with their hypothesized dimensions than with the other dimensions of the instrument.²⁰

There were no floor or ceiling effects for the global K-FOSQ. However, marked ceiling effects were found in all the K-FOSQ subscales except the Mental and Physical Activity Level. These findings are similar to the Norwegian and Swedish versions of FOSQ,^{22,24} showing that all the subscales except Activity Level had ceiling effects. The high ceiling effects mean that these subscales may not differentiate between higher levels of function in these domains.

Correlations between the K-FOSQ and other instruments administered in this study provided evidence for concurrent validity. The K-FOSQ global scores had particularly strong correlations ($r > 0.50$) with ESS and SPI-2, and had medium-sized correlations ($r = 0.40-0.50$) with the total scores from the BDI and SF-36. The Mental and Physical Activity Level domain was the most significantly correlated with all tested instruments. Like our study, ESS has been reported to be well correlated with the all FOSQ subscales and global scores, except the Intimacy and Sexual Relationships domain.^{24,25} In this study, this domain had the lowest correlation coefficients for ESS among the domains of K-FOSQ, too. The five K-FOSQ subscales and the global scale discriminated between patients with ESS scores ≤ 10 and ≥ 11 , which was consistent with the Norwegian version of the FOSQ.²² However, in the Swedish version,²⁴ the differences between high and low ESS-scores groups regarding FOSQ were all statistically significant, except for the Intimacy and Sexual Relationships subscale.

The original global FOSQ and its subscales showed low correlations with SF-36 subscales, suggesting that the FOSQ did not have concurrent validity.⁶ In the present study, however, the global K-FOSQ and its subscales were found to have medium-sized correlation with all the SF-36 subscales. This was consistent with the Swedish version of the FOSQ,²⁴ which showed a positive and statistically significant correlation between the global FOSQ and its subscales and all the subscales in the SF-36.

In the present study, we did not find any statistically significant difference in all K-FOSQ scores among different severities of OSA, except for a trend to have lower scores in severe OSA. In the Mental and Physical Activity Level subscale, the severe OSA subgroup had lower scores than the mild/moderate subgroup. This finding was consistent with that from the Thai FOSQ.²⁵ We also did not find any statistically significant relations between the AHI and K-FOSQ global score and subscales, except a trend toward a weak correlation with General Productivity/Social Outcome and Vigilance.

This study has several limitations that should be noted. First,

this study did not include a normal control group. Therefore we did not confirm whether the K-FOSQ could discriminate between normal healthy control and patient groups. Second, the design of the study was cross-sectional. This design did not allow us to estimate some important aspects of reliability and validity, including responsiveness to changes such as continuous positive airway pressure treatment. Third, the participant group was all patients who visited a sleep laboratory for an evaluation of suspected OSA in a tertiary hospital; therefore, the extent to which the results may be generalized is limited.

In conclusion, we developed a K-FOSQ, and examined its reliability and validity in patients who suffered from simple snoring or OSA. The results from this study provide evidence that the K-FOSQ has internal consistency, test-retest reliability, and construct and concurrent validity. The lowest scores were found in the subscale of Mental and Physical Activity Level, which was the most significantly correlated with all tested instruments. In addition, K-FOSQ was found to appropriately differentiate between the patients with and without daytime sleepiness.

Conflicts of Interest

The authors have no financial conflicts of interest.

REFERENCES

- Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;328:1230-5.
- Kang K, Seo JG, Seo SH, Park KS, Lee HW. Prevalence and related factors for high-risk of obstructive sleep apnea in a large Korean population: results of a questionnaire-based study. *J Clin Neurol* 2014;10:42-9.
- Bixler EO, Vgontzas AN, Lin HM, Calhoun SL, Vela-Bueno A, Kales A. Excessive daytime sleepiness in a general population sample: the role of sleep apnea, age, obesity, diabetes, and depression. *J Clin Endocrinol Metab* 2005;90:4510-5.
- Day R, Gerhardstein R, Lumley A, Roth T, Rosenthal L. The behavioral morbidity of obstructive sleep apnea. *Prog Cardiovasc Dis* 1999;41:341-54.
- Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991;14:540-5.
- Weaver TE, Laizner AM, Evans LK, Maislin G, Chugh DK, Lyon K, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep* 1997;20:835-43.
- American Academy of Sleep Medicine, Iber C. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. Westchester, IL: American Academy of Sleep Medicine 2007.
- Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36 health survey: manual and interpretation guide*. 2nd ed. Boston, MA: The Health Institute, New England Medical Center 1993.
- Han CW, Lee EJ, Iwaya T, Kataoka H, Kohzuki M. Development of the Korean version of Short-Form 36-Item Health Survey: health related QOL of healthy elderly people and elderly patients in Korea. *Tohoku J Exp Med* 2004;203:189-94.
- Cho YW, Lee JH, Son HK, Lee SH, Shin C, Johns MW. The reliability and validity of the Korean version of the Epworth sleepiness scale. *Sleep Breath* 2011;15:377-84.
- Hays RD, Stewart AL. Sleep measures. In: Stewart AL, Ware JE Jr. *Measuring functioning and well-being: the medical outcomes study approach*. Durham; London: Duke University Press 1992.
- Hays RD, Martin SA, Sesti AM, Spritzer KL. Psychometric properties of the Medical Outcomes Study Sleep measure. *Sleep Med* 2005;6:41-4.
- Kim MK, You JA, Lee JH, Lee SA. The reliability and validity of the Korean version of the Medical Outcomes Study-sleep scale in patients with obstructive sleep apnea. *Sleep Med Res* 2011;2:89-95.
- Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4:561-71.
- Han HM, Yum TH, Shin YW, Kim KH, Yoon DJ, Jung KJ. A standardization study of Beck Depression Inventory in Korea. *J Korean Neuro-psychiatr Assoc* 1986;25:487-500.
- Tabachnick BG, Fidell LS. *Using multivariate statistics*. Boston, MA: Allyn and Bacon 2001.
- Norman GR, Streiner DL. *Biostatistics: the bare essentials*. 2nd ed. Hamilton, ON: B.C. Decker 2000.
- Nunnally JC, Bernstein IH. *Psychometric theory*. New York: McGraw-Hill 1994.
- Stevens J. *Applied multivariate statistics for the social sciences*. Hillsdale, NJ: Erlbaum Associates 1986.
- Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 1959;56:81-105.
- Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34-42.
- Stavem K, Kjelsberg FN, Ruud EA. Reliability and validity of the Norwegian version of the Functional Outcomes of Sleep Questionnaire. *Qual Life Res* 2004;13:541-9.
- Vidal S, Ferrer M, Masuet C, Somoza M, Martínez Ballarín JI, Monasterio C. [Spanish version of the Functional Outcomes of Sleep Questionnaire: scores of healthy individuals and of patients with sleep apnea-hypopnea syndrome]. *Arch Bronconeumol* 2007;43:256-61.
- Korpe L, Lundgren J, Dahlström L. Psychometric evaluation of a Swedish version of the Functional Outcomes of Sleep Questionnaire, FOSQ. *Acta Odontol Scand* 2013;71:1077-84.
- Banhiran W, Assanasen P, Methetrairut C, Nopmaneejumrulers C, Chotinaiwattarakul W, Kerdnoppakhun J. Functional outcomes of sleep in Thai patients with obstructive sleep-disordered breathing. *Sleep Breath* 2012;16:663-75.
- Johnson TP, Cho YI, Holbrook AL, O'Rourke D, Warnecke RB, Chavez N. Cultural variability in the effects of question design features on respondent comprehension of health surveys. *Ann Epidemiol* 2006;16:661-8.

- APPENDIX -

K-FOSQ

	(0) 나는 다른 이유로 이 활동을 하지 않는다	(4) 어려움 없음	(3) 예, 약간 어려움	(2) 예, 중등도 어려움	(1) 예, 많이 어려움
1. 당신은 졸리거나 피곤하기 때문에 당신이 하는 일에 집중하기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. 당신은 졸리거나 피곤하기 때문에 일들을 기억하기가 일반적으로 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. 당신은 졸게되거나 피곤해지기 때문에 식사를 끝마치기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. 당신은 졸리거나 피곤하기 때문에 취미활동을 계속하기가 어렵습니까? (예를 들면, 바느질, 수집, 정원가꾸기)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. 당신은 졸게되거나 피곤해지기 때문에 (160 km보다) 짧은 거리를 자동차 운전하기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. 당신은 졸게되거나 피곤해지기 때문에 (160 km보다) 긴 거리를 자동차 운전하기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. 당신은 운전하거나 대중교통을 이용하는 데 너무 졸리거나 피곤하기 때문에 일을 다 마치기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. 당신은 졸리거나 피곤하기 때문에 재정업무를 처리하거나 서류작업을 하기가 어렵습니까? (예를 들면, 수표쓰기, 청구서 지불하기, 재무기록 유지하기, 납세신고용지 적어넣기 등)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. 당신은 졸리거나 피곤하기 때문에 취직일 또는 자원봉사일을 하기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. 당신은 졸게되거나 피곤해지기 때문에 전화대화를 유지하기 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. 당신은 졸게되거나 피곤해지기 때문에 <u>당신집에서</u> 가족이나 친구들과 이야기하기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. 당신은 졸게되거나 피곤해지기 때문에 가족이나 친구들과 <u>그들의 집에서</u> 이야기하기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. 당신은 너무 졸리거나 피곤하기 때문에 당신의 가족이나 친구들을 위해 일을 하기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. 당신은 졸게되거나 피곤해지기 때문에 영화나 비디오테이프를 보기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. 당신은 졸게되거나 피곤해지기 때문에 연극이나 강의를 즐기기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. 당신은 졸게되거나 피곤해지기 때문에 음악회를 즐기기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. 당신은 졸리거나 피곤하기 때문에 텔레비전을 보기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. 당신은 졸리거나 피곤하기 때문에 종교적봉사, 회의, 또는 단체 또는 클럽에 참여하기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. 당신은 졸리거나 피곤하기 때문에 아침에 당신이 원하는 만큼 활동적이기가 어렵습니까?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Korean Version of Functional Outcomes of Sleep Questionnaire

(1) 매우 낮게	(2) 낮게	(3) 중간 높게	(4) 높게
--------------	-----------	--------------	-----------

20. 당신은 당신 활동의 일반적 수준을 어떻게 평가하시겠습니까?

(0) 나는 다른 이유로 이 활동을 하지 않는다	(4) 어려움 없음	(3) 예, 약간 어려움	(2) 예, 중등도 어려움	(1) 예, 많이 어려움
--	------------------	---------------------	----------------------	---------------------

21. 당신은 졸리거나 피곤하기 때문에 이성과의 깊은 관계 또는
성관계가 지장을 받습니까?

22. 당신은 졸리거나 피곤하기 때문에 이성과의 깊은 사귀에 대한
욕구 또는 성에 대한 욕구가 지장을 받습니까?

23. 당신은 졸리거나 피곤하기 때문에 성적으로 흥분하게 되는
능력이 지장을 받습니까?

24. 당신은 졸리거나 피곤하기 때문에 성적흥분의
최고조(오르가즘)를 갖는 능력이 지장을 받습니까?