



# Psychometric Properties of the Arabic Version of the Pittsburgh Sleep Quality Index in Hemodialysis Patients

Ahmed N. Albatineh<sup>1</sup> · Abdullah Al-Taiar<sup>2</sup> · Reem Al-Sabah<sup>1</sup> · Bashar Zogheib<sup>3</sup>

Received: 24 March 2022 / Revised: 19 June 2022 / Accepted: 3 July 2022 / Published online: 27 July 2022  
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd. 2022

## Abstract

**Purpose** The Pittsburgh Sleep Quality Index (PSQI) is widely used in screening for self-reported sleep disturbances in clinical groups or healthy populations. The literature lacks rigorous psychometric evaluation of the Arabic Pittsburgh Sleep Quality Index (A-PSQI) in hemodialysis (HD) patients. This study aims to estimate reliability, extract, and test the underlying factor structure of the A-PSQI in HD patients, and determine its suitability for this clinical group.

**Methods** In a sample of 461 HD patients (287 males, 62.3%) recruited from all health districts in Kuwait., inter-components, between A-PSQI components, and its global score correlations were estimated. Component analysis was conducted to identify inconsistent components with the rest. Cronbach's  $\alpha$  and McDonald's  $\omega$  estimated reliability. Stability and convergent validity were assessed. Construct validity was explored by exploratory factor analysis (EFA). Confirmatory factor analysis (CFA) was conducted to test factor structure.

**Results** Patients had a median (Interquartile Range) age of 54 (22) years. The A-PSQI had a mean (SD) score of 7.1 (3.65). The Cronbach  $\alpha = .634$  and McDonald  $\omega = 0.62$  indicated moderate reliability. EFA resulted in two factors highly correlated (0.83,  $PV < 0.001$ ) with A-PSQI global score indicating two distinct constructs. CFA provided weak evidence supporting the one-factor model, but the two-factor model (explained 51.8% of total variance) without the "use of sleep medication" item, which has the smallest corrected-item total correlation, showed a good fit with the smallest AIC.

**Conclusion** The A-PSQI seems moderately reliable to screen for sleep disorders in HD patients. Future research on scoring in the two-factors model is warranted. It can help clinicians better understand sleep disturbance and improve the instrument's sensitivity.

**Keywords** PSQI · Hemodialysis · Reliability · EFA · CFA · Arabic

## 1 Introduction

Quality of sleep is an important aspect of a person's life because it contributes to enhanced functioning, psychological well-being, and improved quality of life [1]. Several diseases are associated with poor sleep including but

not limited to cancer [2, 3], heart failure [4], schizophrenia [5], depression, and generalized anxiety disorder [6]. Lack of sleep may lead to reduced cognitive functioning [7] and low functional cell-mediated immunity response [8]. Poor quality of sleep is a major health issue among Hemodialysis (HD) patients with a prevalence ranging from 41 to 83% [9]. The presence of poor quality of sleep in HD patients has been associated with reduced quality of life [10] and increased mortality [11]. Moreover, poor quality of sleep has been reported to be involved in the development of cardiovascular diseases in patients undergoing maintenance HD [12].

Quality of sleep can be measured objectively using polysomnography or electroencephalographic spectral component analysis which can provide reliable data. But these are impractical tools in large-scale studies and hence screening instruments become the ideal alternative at least

✉ Ahmed N. Albatineh  
ahmed.albatineh@ku.edu.kw

<sup>1</sup> Department of Community Medicine and Behavioral Sciences, Faculty of Medicine, Kuwait University, P. O. Box 24923, 13110 Safat, Kuwait

<sup>2</sup> School of Community and Environmental Health, College of Health Sciences, Old Dominion University, Norfolk, VA, USA

<sup>3</sup> Department of Mathematics and Natural Sciences, American University of Kuwait, Salmiya, Kuwait

to screen for sleep disorders [13]. The Pittsburgh Sleep Quality Index (PSQI) is a widely used subjective self-report screening instrument developed in 1989 as a one-factor structure with a global score representing the quality of sleep [14]. It has been used in the general population and among several clinical groups [1, 13, 15]. The PSQI has been established as a valid instrument with acceptable psychometric properties in the adult population [15], pregnant women [16], cancer patients receiving chemotherapy [1], patients with sleep apnea syndrome, depressed patients, and individuals with dementia [17], coronary artery disease patients [18], patients with posttraumatic stress disorder [19], and renal transplant recipients [20]. The PSQI has been translated into several languages and was used in ethnically different populations, including Greek [1], Italian [17], Japanese [5], Spanish [21], Portuguese [22], Ethiopian (Amharic)[15], and Korean [23].

While the PSQI was psychometrically evaluated in several non-clinical and clinical populations, there is a paucity of research that evaluated the reliability, factor structure, and construct validity of PSQI among HD patients. In that regard, there was only one research paper aimed to investigate the factor structure of the PSQI among renal transplant recipients [20] in which a three-factor structure was tested. In an extensive literature search for validation of PSQI among Arabic speaking patients, a recent study has attempted to validate PSQI in 369 cancer patients [24]. This study implemented Confirmatory Factors Analysis (CFA) and concluded that a refined one-factor structure showed an acceptable fit for cancer patients. In another study, the PSQI was validated among 130 Arabic speaking patients with coronary artery disease [18]. In that study, PSQI was tested against the insomnia severity index (ISI) and the medical outcome study short form 36 (SF-36) using Pearson correlation, but neither the factor structure nor the construct validity of PSQI using CFA were investigated. In a descriptive co-relational study, the PSQI was tested in a sample of 35 healthy Arabic bilinguals with ISI as the retest instrument using Pearson correlation but without exploring its factor structure and construct validity [18]. Finally, in a systematic review of psychometric properties and cultural adaptation of four sleep disturbance instruments used in Arabic speaking populations, including PSQI; Al Maqbali et al. [24] recommended further psychometric testing and cultural adaptation for these scales.

Extensive literature search revealed that to date, no previous study has validated A-PSQI in patients on HD treatment and its factor structure and construct validity remain unknown. Therefore, the main goals of this study are (1) test the reliability of the A-PSQI among HD patients (2) use EFA to investigate the factor structure of the A-PSQI (3) use CFA to test its construct validity by testing the factor structure using several goodness-of-fit indices.

## 2 Materials and Methods

### 2.1 Study Design and Data Collection

This is a cross-sectional study conducted with the aim to translate the PSQI instrument into Arabic language, test its reliability and factor structure in HD patients recruited from all health districts in Kuwait. Data were collected from all six governmental dialysis centers in Kuwait treating more than 1800 patients using a non-probabilistic sample of 370 HD patients. Patients above 18 years, mentally and clinically stable, and clearly understand Arabic language were included in the study. Patients who have hearing impairments or mental health problems were excluded. Patients were interviewed face-to-face and informed consents and ethical approval by ethics committee at Kuwait Ministry of Health (MOH) was obtained.

### 2.2 Arabic Version of PSQI

The PSQI is a widely used self-report instrument in screening for sleep quality over one-month interval [14]. The PSQI scale consists of 11 questions of 19 individual items, and each item is weighted on 0–3 scale; these items create 7 components during the analysis, which produce a total score ranging from 0 to 21 [14]. A total score of 5 or more indicates a bad sleep, while a score lower than 5 indicates good sleep [14]. PSQI is currently the only instrument that covers a broad range of indicators relevant to sleep quality and has been described as a reliable and valid instrument [25]. The PSQI was translated into Arabic language by two independent researchers and then back translated into English by two other independent researchers and any conflicts were resolved through discussion. The tool was pilot tested on 15 patients for clarity whose responses were discarded. The A-PSQI was further checked for clarity and suitability by a psychologist (one of the authors) and a physician working at one of the visited dialysis centers. Their comments were integrated in the final A-PSQI version.

### 2.3 Statistical Analysis

Statistical analyses were conducted using SPSS (IBM version 25) and R [26] statistical software. Since factor analysis assumes that the variables are correlated to some degree, and since PSQI components were measured on a Likert scale, the more appropriate polychoric correlation was used to estimate correlations between the seven components by the “polycor” R package. The Kaiser Meyer Olkin (KMO) index and the Bartlett test of sphericity (BTS) were calculated

prior to factor extraction to determine sample size adequacy and appropriateness of the correlation matrix, i.e., significant correlations ( $PV < 0.05$ ) among at least some of the A-PSQI components. Values of KMO closer to one indicate better adequacy of the sample size to perform factor analysis. KMO values between 0.6 and 0.69 are acceptable, between 0.7 and 0.79 are considered good, between 0.8 and 0.9 are very good, while  $> 0.9$  is excellent [27]. Ceiling or floor effects were present if  $> 15\%$  of HD patients achieved the highest or lowest A-PSQI score, respectively [28].

Internal consistency indicates the extent to which the components in A-PSQI measure the same construct. This helps the researcher to determine the components that are not consistent with the test in measuring quality of sleep and remove them in order to improve internal consistency and hence higher chances of A-PSQI being reliable [29]. Reliability analysis was conducted using Cronbach's  $\alpha$  [30] and McDonald  $\Omega$  [31]. A Cronbach's  $\alpha > 0.70$  is considered acceptable for an instrument's internal reliability [32]. McDonald  $\Omega$  is an estimate of the general factor saturation of a test and is considered by some researchers the best estimate for reliability [33]. In addition to Cronbach's  $\alpha$  and McDonald  $\Omega$ , item analysis was conducted. Item analysis provides a refinement of the test reliability by identifying "problem" items in the test, i.e., items that yield low correlations and removing them will improve reliability [29]. Furthermore, reliability was estimated using split-half method; an alternative to test-retest reliability, which measures the extent to which all items of the test contribute equally to the constructs being measured. Spearman-Brown coefficient and Guttman split-half coefficient were estimated [34, 35]. The split-half method is ideal to reduce bias that might be introduced during the retest phase due to respondents' unwillingness or inability to take the survey a second time. Finally, the face validity of the instrument was tested on 10 individuals with diverse educational levels to assess comprehension and acceptability by HD patients. Finally, one question related to sleep was used to assess convergent validity. For sleep ability, "I can lie-down without problems" or "I must raise bed-head to breath" coded as 0 or 1 was correlated with A-PSQI global score by point-biserial correlation. HD patients who "must raise bedhead to breath" are expected to score higher on A-PSQI global score (bad sleepers) and the correlation to be "medium" according to guidelines (Cohen 1988).

Exploratory factor analysis (EFA) was conducted using principal components analysis (PCA) applied to the correlation matrix of the A-PSQI seven components. Extracted factors that accounted for large portion of the variability were determined based on the eigenvalue approach with eigenvalues greater than one being included. Since the extracted factors generally have the possibility of being highly correlated, the oblimin non-orthogonal factor rotation with Kaiser Normalization was implemented in the extraction process to

keep the rotation of axis more flexible to discover any overlapping between the factors and hence produce more accurate clustering of the variables [29]. A factor loading was set at  $\pm 0.33$  or more for practical significance. The justification was that such factor loading represents the amount of the variable's total variance ( $(0.33)^2 = 0.10$ ) accounted for by the factor [29]. This will help generate better interpretations and clustering of the data which is in line with the approach of the investigators who used EFA [36].

Confirmatory factor analysis (CFA) was conducted using "lavaan" R statistical package [37], and the extracted factors using EFA from the A-PSQI data were tested. A hypothesized one-factor model representing a global A-PSQI score [14] and a two-factor model extracted from the data were both tested. Several goodness-of-fit indices were used to evaluate the proposed models including comparative fit index (CFI) [38], Akaike information criterion (AIC) [39] which allows for comparison between the models, root mean square error of approximation (RMSEA) [40], standardized root mean square residual (SRMSR), and Bentler-Bonnet non-normed fit index (NNFI) [41].  $CFI > 0.90$  indicates acceptable fit, while  $CFI \geq 0.95$  indicates good fit [41].  $RMSEA < 0.08$  indicates acceptable fit [42] while  $RMSEA < 0.05$  indicates good fit [43] and  $NNFI \geq 0.95$  indicates good fit [41]. Finally,  $\chi^2$  goodness-of-fit test was used to determine acceptability of the model fit with a significant  $\chi^2$  implying poor fit model, i.e., significant proportion of variance within the data is unexplained by the model [44]. However, it must be noted that  $\chi^2$  test assumes multivariate normality, sensitive to large sample size, and for small variation in the data,  $\chi^2$  tends to be significant [41]. For these reasons, the ratio  $\chi^2/\text{dof}$  will be used with  $\chi^2/\text{dof} < 3$  indicates acceptable model fit [45]. These model fit indices are less affected by sample size, model misspecification, and parameter estimates [46, 47].

## 3 Results

### 3.1 Descriptive Findings

The patients' age (has skewed distribution) ranged from 21 to 95 years with median 54 and IQR 22 years. The distribution of HD patients according to socio-economic factors and disease profile characteristics is presented in Table 1. Patients were mostly males (62.3%), Kuwaiti nationals (86.6%), married (68.3%), with high school education or less (52.9%), retired (47.3%), with income  $< 1000$  KD (52.5%), had diabetes as primary cause of dialysis (45.3%), non-smokers (78.3%), can lie down on bed without problems (68.5%). The prevalence of sleep disorder among HD patients was 66.3% using the cut-off point  $\geq 5$  for PSQI [14]. Overall, the total A-PSQI score did not have floor or ceiling

**Table 1** Distribution of HD patients according to socio-economic factors, disease profile characteristics in Kuwait, February 2018 (N=461)

Characteristic	n <sup>1</sup> (%)
<i>Gender</i>	
Male	287 (62.3)
Female	174 (37.7)
<i>Marital status</i>	
Single	65 (14.1)
Married	315 (68.3)
Divorced	43 (9.3)
Widowed	38 (8.2)
<i>Dialysis duration (Yrs.)</i>	
<2	41 (8.9)
2–4	389 (84.4)
>4	31 (6.7)
<i>Primary cause of dialysis</i>	
Diabetes	209 (45.3)
Hypertension	116 (25.2)
Glomerulonephritis	60 (13.0)
PCKD	34 (7.4)
Other	37 (8.0)
<i>Sleep ability</i>	
I can lie down without problems	316 (68.5)
I must raise bed head to breath comfortably	145 (31.5)
<i>Education</i>	
High school or below	244 (52.9)
Diploma	109 (23.6)
University degree	96 (20.8)
Professional/Higher education	11 (2.4)
<i>Nationality</i>	
Kuwaiti	399 (86.6)
Non-Kuwaiti	62 (13.4)
<i>Employment</i>	
Student	18 (3.9)
Employed	120 (26.0)
Unemployed	104 (22.6)
Retired	218 (47.3)
<i>Family Income</i>	
Less than 500 KD	119 (25.8)
500 to 999 KD	123 (26.7)
1000 to 1500 KD	117 (25.4)
Over 1500 KD	102 (22.1)
<i>Smoking status</i>	
Smoker	100 (21.7)
Non-smoker	361 (78.3)

<sup>1</sup>Some values are missing, PCKD Polycystic Kidney Disease, KD Kuwaiti Dinar

effects; 0.9% of HD patients reported a minimum score of zero, and none reported a maximum score of 21.

### 3.2 Internal Consistency

Reliability as measured by internal consistency using Cronbach  $\alpha$  based on standardized items for A-PSQI (all components) was 0.634. Assuming a two-factor structure exists, the Cronbach  $\alpha$  for factor 1 (components 1,2,5,6,7) was 0.630 and 0.562 for factor 2 (components 2,3,4). Furthermore, according to McDonald  $\omega$ , the estimated reliability for the PSQI (all components) was 0.623 and with two-factor structure assumed, the McDonald  $\omega$  was 0.622 and 0.628 for factor 1 and factor 2, respectively. According to split-half method for all A-PSQI components, part one included four components (components 1,2,3,4) had Cronbach  $\alpha$ =0.56, while part two included 3 components (components 5,6,7) had Cronbach  $\alpha$ =0.50 with Spearman-Brown coefficient 0.483 and Guttman Split-Half Coefficient 0.458. The reliability is further supported by the moderate to high correlations between each of the seven components and the global A-PSQI score which ranged between 0.58 and 0.70. Inter-components correlations were estimated using polychoric correlations. Internal homogeneity was estimated using Pearson correlations between A-PSQI components and the global A-PSQI score, and inter-components correlations are presented in Table 2. To refine the reliability by identifying components in A-PSQI that yield low correlations with the sum of the scores on the remaining items, item analysis was conducted to increase internal consistency of A-PSQI by removing components with low correlations. Components were retained if their item-total correlation is at least 0.30 which indicates that about 10% of the variance in A-PSQI is accounted for by that item [29]. Corrected item-total correlations along with individual items reliability statistics for A-PSQI are presented in Table 3. Values of the corrected item-total correlations indicated that component 6 (sleeping medication use) can be dropped. In fact, the results in Table 3 showed that, compared to other components, if component 6 was deleted, Cronbach  $\alpha$  is closer to the value when component 6 was present. Also, it is worth noting that component 6 is poorly correlated with components 2, 3, and 4 as presented in Table 2.

For convergent validity, A-PSQI mean score was significantly higher among those “can lie-down without problems” (mean = 6.5, SD = 3.5) compared to those “must raise bed-head to breath” (mean = 8.4, SD = 3.7). As expected, the correlation of 0.232 (PV < 0.001) (95% CI: 0.144, 0.317) was significant of medium magnitude.

### 3.3 Exploratory Factor Analysis

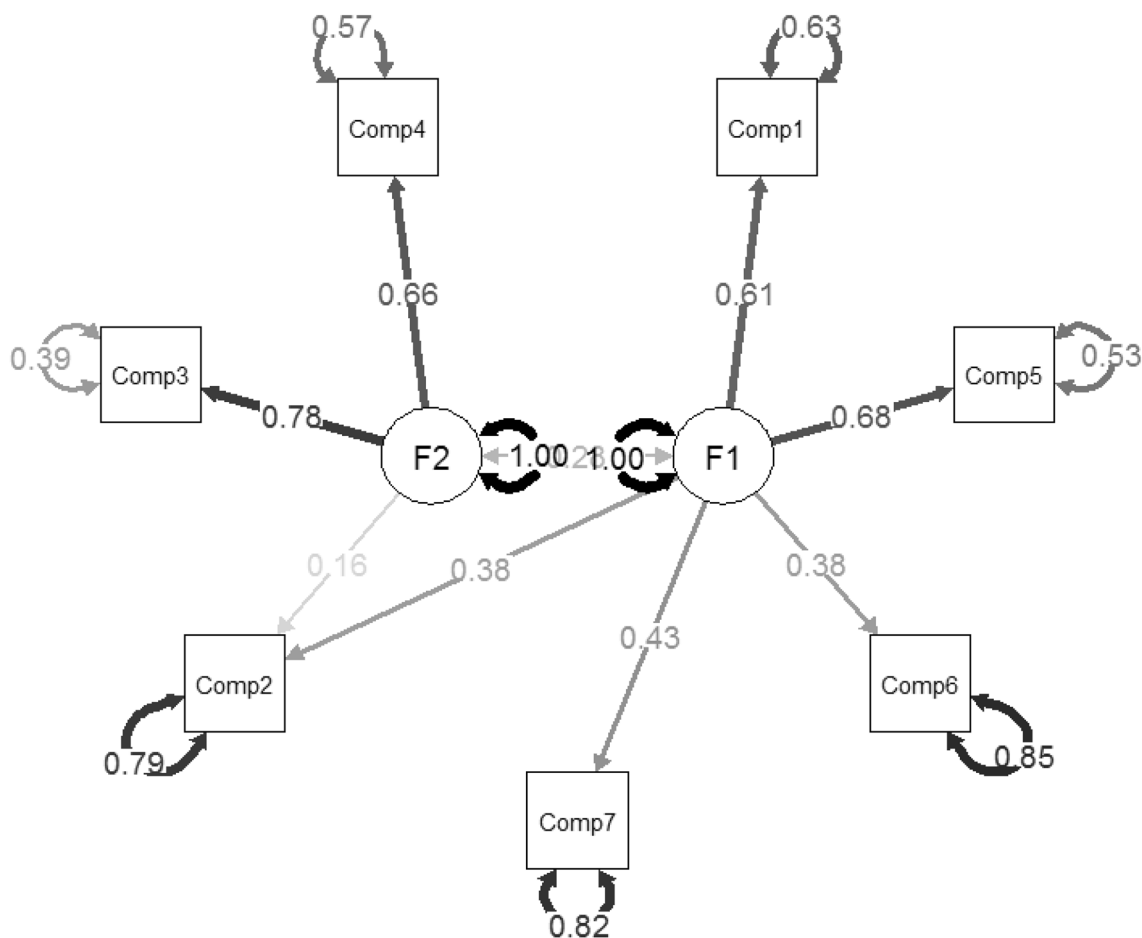
The KMO measure of sampling adequacy for all A-PSQI components was 0.68, while the sampling adequacy measures for individual components; given by the anti-image

correlations, were all above 0.70, except for components 3 and 4 which were around 0.6 (Table 3). The BTS had  $\chi^2=423.1$ , (dof=21,  $PV < 0.001$ ). This indicates that factor analysis is warranted. Following PCA factor extraction and oblimin non-orthogonal rotation with Kaiser Normalization, two factors with eigenvalues  $> 1$  ( $\lambda_1 = 2.23$ ,  $\lambda_2 = 1.39$ , and %variance explained = 31.9%, 19.9%, respectively) were extracted using all seven A-PSQI components. The two factors accounted for 51.8% of the total variance. Factor 1 represents the components “subjective sleep quality”, “sleep latency”, “sleep disturbances”, “sleeping medication use”, and “daytime dysfunction” while factor 2 represents the components “sleep duration”, “habitual sleep efficiency”, and “sleep latency” which has cross-loading on both factors with all components’ factor loadings presented in Table 3. For each extracted factor, the factor scores were calculated for each HD patient by regression method. According to Martin et al. [48], unlike the method proposed by Anderson-Rubin of factor scores calculations, regression method does not assume the

extracted factors to be orthogonal and also minimizes any sum of squares discrepancies between the true and estimated factors over individuals. In that regard, the Pearson correlation coefficient between factor 1 and factor 2 was 0.166 ( $PV < 0.001$ ). The low correlation between the two factors may indicate that the two factors are representing two different constructs and hence a two-factor model can be a good fit. Furthermore, Pearson correlation between A-PSQI global scores and factor scores generated using regression method were 0.83 ( $PV < 0.001$ ) for factor 1 and 0.68 ( $PV < 0.001$ ) for factor 2; another indication that each factor represent a separate construct.

### 3.4 Confirmatory Factor Analysis

Several models were tested, and their goodness-of-fit indices are reported in Table 4. The ratio  $\chi^2/\text{dof} < 3$  was attained by models 2, 3, 4 and so potentially may indicate good fit. This means that significant portion of the variation within the data is accounted (explained) for by these models [44].



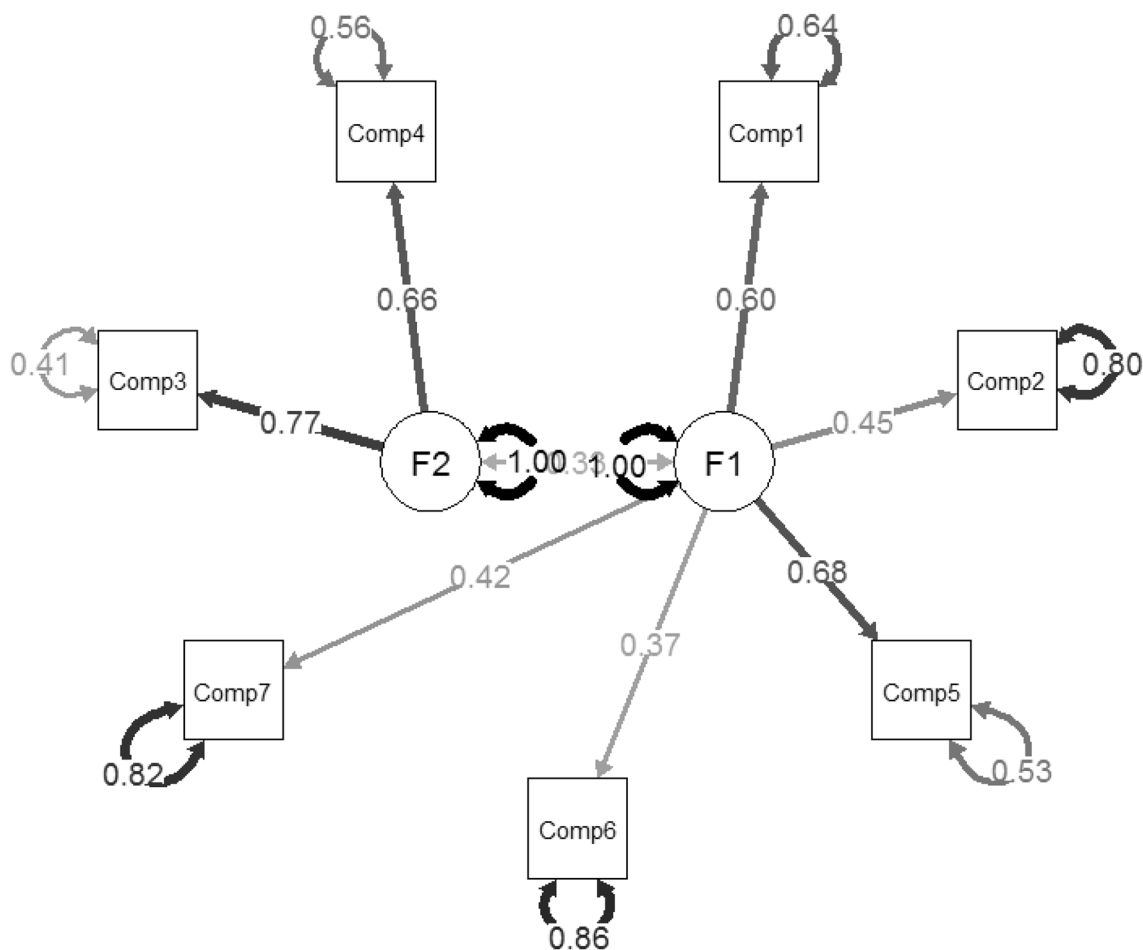
**Fig. 1** Path diagram for the two-factor model (model 2 in Table 4) extracted using PCA for PSQI scale applied to HD patients in Kuwait (N=448)

According to Hu and Bentler [41], for model evaluation, we will furthermore resort to CFI, NNFI, RMSEA, and SRMSR to determine goodness-of-fit and AIC for models' comparison [43].

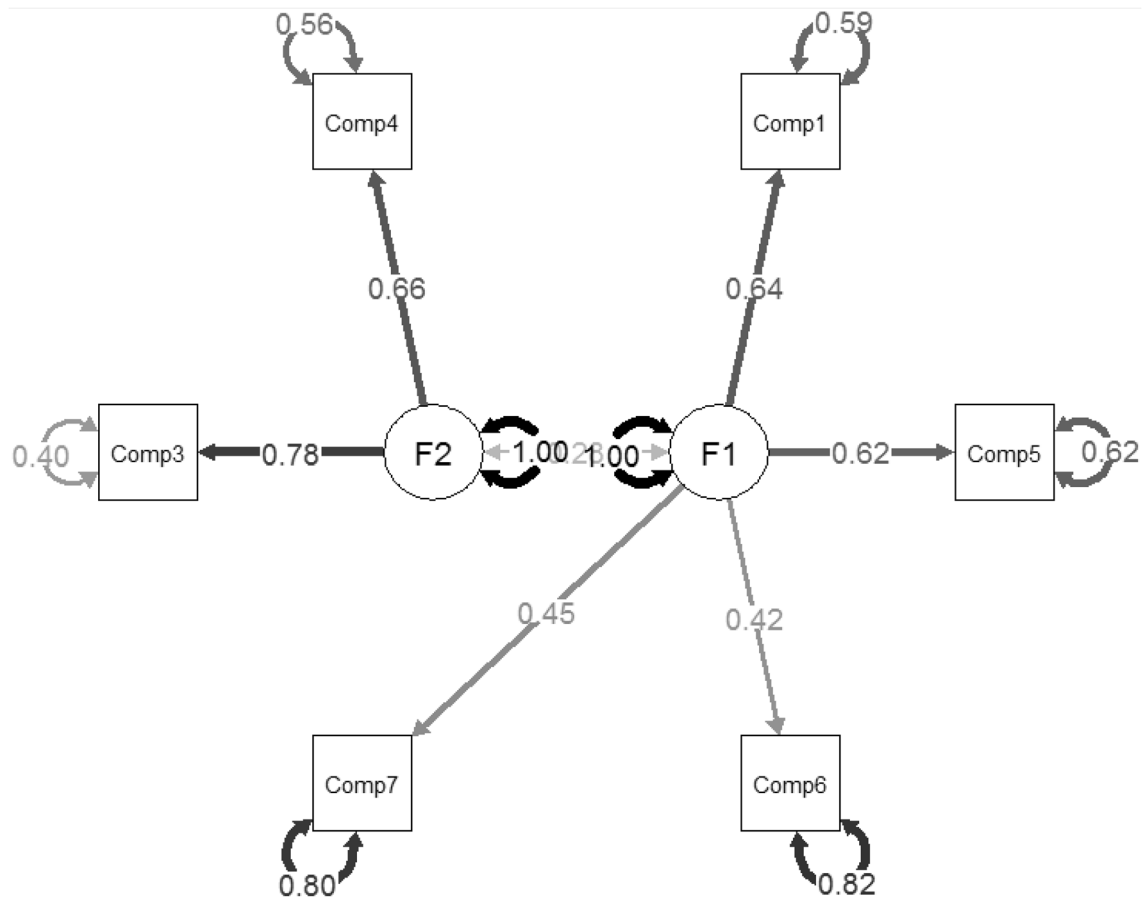
Results indicated that the one-factor model (model 1 in Table 4) composed of all seven components has poor fit according to all goodness-of-fit indices. Furthermore, for the reasons discussed earlier, component 6 was dropped and a six-component model was fitted and showed poor fit (results not shown). This clearly indicates that sleep disturbance is not a one-dimensional construct in HD patients. On the other hand, the two-factor model, as presented in Fig. 1, which accounted for 51.8% of the total variance (model 2 in Table 4) factor 1 (components 1,2,5,6,7) and factor 2 (components 2,3,4) extracted by PCA with cross-loading allowed on both factors was fitted and results indicated it has good fit indices except for NNFI with border value. If cross-loading was allowed on the factor with higher loading [29], and the two-factor model was

fitted as presented in Fig. 2 (see model 3 of Table 4), then according to all goodness-of-fit indices it has good fit except for NNFI. The last option to deal with a component with cross-loading is to delete it [29] and if this is applied and fitted (model 4 in Table 4) as presented in Fig. 3, then this model has good fit according to all goodness-of-fit indices as shown in Table 4. Finally, the Q-Q plots of the standardized residuals of the A-PSQI data from HD patients obtained using correlation matrices from models 1, 2, 3 and 4 in Table 4 fitted using CFA are presented in Fig. 4. It is worth noting that for comparison purposes, model 4 has the smallest AIC and hence the model with best performance.

An attempt was made to extract three factors from the HD data by considering factors corresponding to the largest three eigenvalues ( $\lambda_1 = 2.231$ ,  $\lambda_2 = 1.393$ ,  $\lambda_3 = 0.915$ ). Although the total variance explained was 64.8% (compared to 51.8% for two-factors), but CFA showed that this three-factor model had poor fit indices (results not shown).



**Fig. 2** Path diagram for the two-factor model (model 3 in Table 4) extracted using PCA for PSQI scale applied to HD patients in Kuwait ( $N=448$ )



**Fig. 3** Path diagram for the two-factor model (model 4 in Table 4) extracted using PCA for PSQI scale applied to HD patients in Kuwait ( $N=448$ )

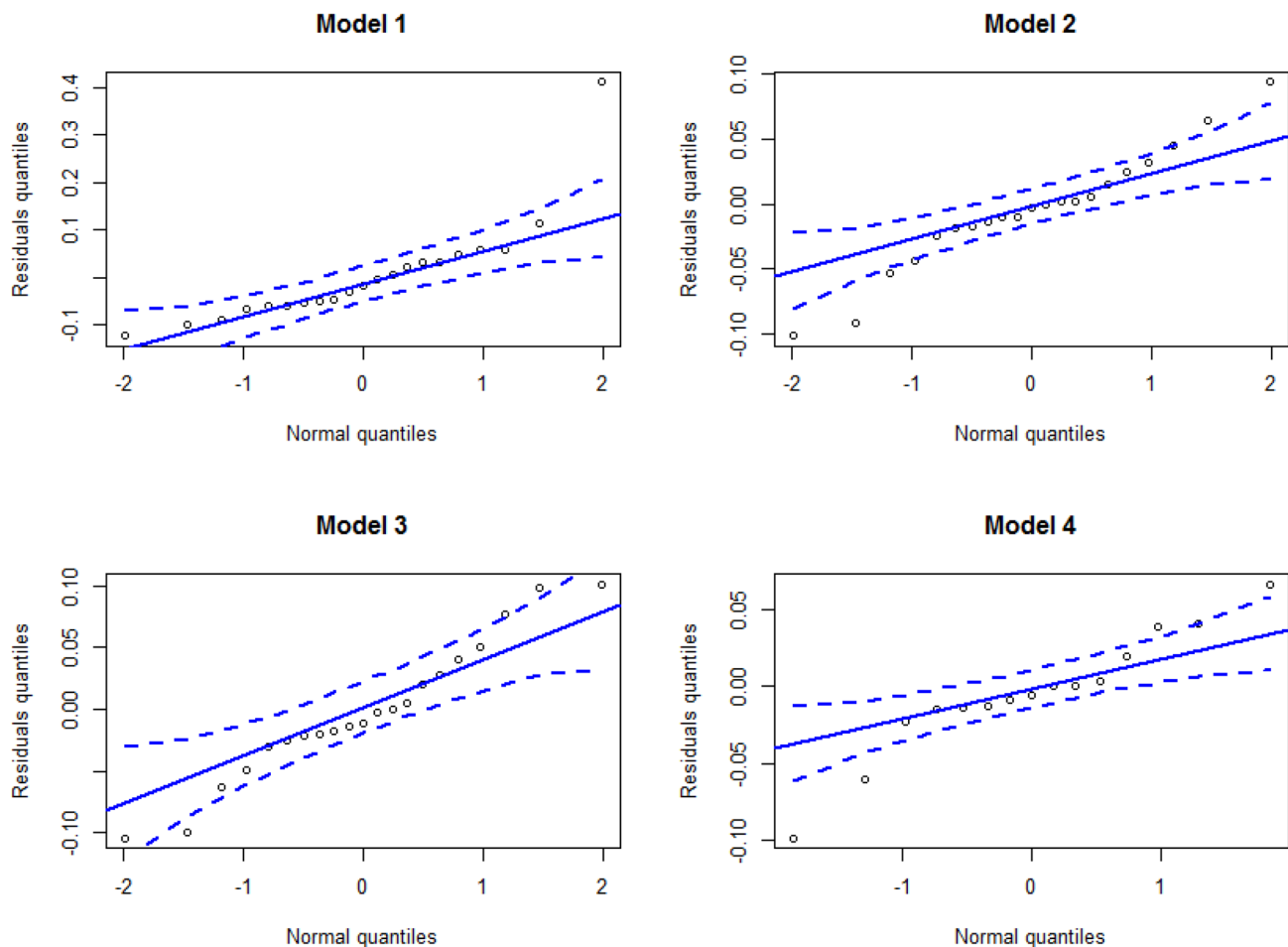
## 4 Discussion

This is the first study to explore the factor structure and construct validity of A-PSQI among HD patients. Moreover, to the best of our knowledge, the factor structure and construct validity of PSQI has not been explored among English speaking HD patients, peritoneal dialysis (PD) patients, but was explored among renal transplant recipients [20]. Several studies have investigated the psychometric properties of PSQI either using a gold standard instrument or a clinical test to estimate PSQI sensitivity and specificity to assess sleep quality [17, 18, 23, 49, 50]. Other studies investigated the factor structure and construct validity using EFA and CFA [1, 15, 16, 20, 24, 51, 52]. In this study, we used PCA to extract latent variables by an oblimin rotation and test its validity using CFA using several goodness-of-fit indices.

The A-PSQI showed moderate internal consistency as measured by Cronbach  $\alpha$  and McDonald  $\omega$ . Although Cronbach  $\alpha$  leaned toward the lower side, but it is still comparable and sometimes higher compared with published research

[15, 16, 18, 51]. Cronbach  $\alpha$  can be affected by dimensionality (number of latent factors), number of items in the scale, and average inter-item correlation among the items. Larger number of items and average inter-item correlations can lead to higher values of Cronbach  $\alpha$ , and the converse is true [53]. For this study, as indicated in Table 2, components 6 and 7 had poor correlations with components 2, 3, and 4 which probably caused Cronbach  $\alpha$  to be in its current range.

The original clinical formulation of PSQI by its authors [14] suggested a one-factor model of seven components that form a global score for sleep quality. The results in the current study indicated the one-factor model had poor fit according to all goodness-of-fit indices and hence did not capture the two or even three-dimensional nature of sleep quality measured by A-PSQI. This result is consistent with that reported in renal transplant recipients [20]. Due to insensitivity of Cronbach  $\alpha$  to deletion of component 6 (sleeping medication use) and due to its smallest corrected-item total correlation (Table 3), a one-factor model without component 6 was fitted but produced poor fit indices (results not reported). This led us to believe that the one-factor



**Fig. 4** Q-Q plots of the standardized residuals of the A-PSQI data from HD patients obtained using correlation matrices from models 1, 2, 3 and 4 in Table 4 fitted using CFA ( $N=448$ )

**Table 2** Inter-item correlation matrix for A-PSQI components and global A-PSQI score based on polychoric\* correlations for HD patients in Kuwait ( $N=448$ )

Component	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	PSQI
Comp1	1.00							
Comp2	0.292	1.00						
Comp3	0.204	0.2407	1.00					
Comp4	0.145	0.2238	0.6390	1.00				
Comp5	0.485	0.4283	0.1773	0.2078	1.00			
Comp6	0.360	0.0764	-0.0216	0.0103	0.368	1.00		
Comp7	0.328	0.1675	0.0800	0.0832	0.329	0.3461	1.00	
PSQI	0.698	0.6033	0.6204	0.6328	0.687	0.5866	0.5752	1.00

\*Polychoric correlations are provided for descriptive purposes only and were not analyzed for significance

model for PSQI in Arabic HD patients was inappropriate. This conclusion is consistent with other published research who reported weak evidence in support of one-factor model of a global score for sleep quality [1, 20, 52].

In the current study, the extracted two-factor model, which in essence is consistent with the factor structure from

the Greek cancer patients or Spanish patients with fibromyalgia [1, 21], though their components structure were not quite the same. Furthermore, an attempt was made to extract three factors from our A-PSQI data in HD patients by considering factors corresponding to the three largest eigenvalues. Although the total variance explained was 64.8%,



**Table 3** Item analysis and factors loadings\* of A-PSQI components following PCA with factor extractions using oblimin rotation method with Kaiser Normalization for HD patients (N=448)

PSQI component	PCA		Item analysis			Scale mean (SD)	MSA
	Factor 1	Factor 2	CI-TC	$\alpha^1$	$\omega^2$		
Subjective sleep quality	0.695	0.106	0.421	0.569	0.551	1.12 (1.03)	0.741
Sleep latency	0.410	0.351	0.331	0.603	0.584	1.18 (0.94)	0.744
Sleep duration	0.022	0.836	0.332	0.607	0.602	1.08 (1.10)	0.586
Habitual sleep efficiency	-0.019	0.827	0.318	0.617	0.606	0.54 (0.97)	0.574
Sleep disturbances	0.702	0.142	0.466	0.561	0.539	1.48 (0.66)	0.719
Sleeping medication use	0.652	-0.257	0.223	0.612	0.600	0.55 (0.96)	0.700
Daytime dysfunction	0.649	-0.097	0.300	0.636	0.621	1.20 (0.91)	0.769
PSQI	Min = 0, Max = 19					7.10 (3.65)	0.690

Factor loading more than 0.33 are considered to be significant, (It ensures that such item can explain at least 10% of the variation  $(0.33)^2=0.10$ ). CI-TC: Corrected item-total correlation.  $\alpha^1$ : Cronbach's alpha if component was deleted.  $\omega^2$ : value of McDonald's  $\omega$  if component was deleted. PCA Principal Components Analysis, MSA Results of Kaiser-Meyer-Olkin test of sampling adequacy for individual items and total A-PSQI score

\*Factor loadings are obtained from the pattern matrix

**Table 4** Goodness-of-fit indices using confirmatory factor analysis applied to A-PSQI for HD patients in Kuwait (N=448)

Models	$\chi^2$ (dof) (PV)	$\chi^2$ /dof	RMSEA (PV)	SRMSR	CFI	NNFI	AIC
Model 1	146.9 (14) (0.001)	10.5	0.146 (0.001)	0.094	0.673	0.509	8158.5
Model 2	25.3 (12) (0.013)	2.1	0.050 (0.465)	0.039	0.967	0.943	8040.8
Model 3	32.9 (13) (0.002)	2.5	0.059 (0.257)	0.047	0.951	0.921	8046.5
Model 4	12.1 (8) (0.149)	1.5	0.034 (0.731)	0.033	0.988	0.977	6888.8

RMSEA Root mean square error of approximation, SRMR Standardized root mean square residual, CFI Comparative fit index, NNFI Bentler-Bonett Non-Normed Fit Index, AIC Akaike Information Criterion. Cut-off points: NNFI  $\geq 0.95$  good fit, CFI  $\geq 0.90$  acceptable fit, CFI  $\geq 0.95$  good fit, RMSEA  $< 0.08$  acceptable fit, RMSEA  $< 0.05$  good fit, SRMR  $< 0.08$  acceptable fit, SRMR  $< 0.05$  good fit. Model 1: One factor model in which all 7 A-PSQI components load on one factor; Model 2: Two-factors' model extracted using PCA allowing for cross-loading; Model 3: Two-factors' model extracted using PCA allowing cross-loading on higher factor only. Model 4: Two-factors' model extracted using PCA with the cross-loading component deleted

but CFA showed that such model produced poor fit indices. However, an attempt was made to fit three-factor model as described by Cole et al. [52], which produced non-positive definite matrix of latent variables.

A three-factor model was reported in other populations including renal transplant recipients [20], Nigerian university students [54], and Peruvian pregnant women [16], but not in our study. It is worth noting that the correlated three-factor model reported in the Peruvian pregnant women showed poor fit indices for our A-PSQI data in HD patients (results not shown in Table 4), however its two-factor model without component 6 (sleeping medication use) showed better fit indices compared to the model with component 6. This clearly indicates that incorporating component 6 weakens the model fit and instrument reliability; a phenomenon worth noting which has been reported by others [15]. A recent systematic review concluded that “sleeping medication use” has been shown to contribute poorly to construct validity of the PSQI [25]. Moreover, in the Peruvian pregnant women study, it was

noted that removal of the “sleeping medication use” component neither improved the fit of the CFA models nor had a noticeable influence on the construct validity of the PSQI [16]. A possible explanation for the insensitivity of Cronbach  $\alpha$  for deletion of “sleeping medication use”, not improving the CFA models, or having any influence on the construct validity is that this component is not measuring the same construct as other components. To avoid missing the factor structure in the A-PSQI for HD patients, we used unweighted least squares rotation method, which is distribution free that yields consistent estimates [55], along with oblimin rotation and found identical factor structure from both methods. This was reassuring that the factor structure obtained is not confounded by the rotation methods, hence made our findings robust against choice of rotation method. Finally, the “sleep latency component” cross-loaded on both factors; although several rotation methods were used, which may indicate that the concept/wording of this component were not clear and so HD patients got confused causing cross-loading. This

study has some limitations including the cross-sectional nature, use of convenient sample, and self-report data that may have introduced some recall bias. Furthermore, sleep diaries of the patients could have been used and then correlated with the A-PSQI score. Some of the strengths of this study include the large sample size which comprised one-fourth of HD patients in Kuwait collected from all six governmental dialysis centers by attending all three shifts on several days.

## 5 Conclusions

This is the first study that investigated the reliability, factor structure, and construct validity of the A-PSQI among HD patients. The results indicated that the A-PSQI is a moderately reliable tool to screen for sleep disorder among HD patients. However, empirical evidence for its use as a global score for screening sleep disturbance in HD patient is not warranted. In this study, two factors were extracted using PCA with “sleep latency” component cross-loading on both factors. The extracted two-factor model without “sleep latency” showed very good fit indices and smallest AIC compared to the models with this component loading on the factor with higher loading or cross loading on both factors. Future research on scoring based on two-factor model in HD patients is warranted. Further confirmatory research of sleep disorder among HD patients is needed. This will hopefully help clinicians identify constructs of sleep disturbance in HD patients more discretely and improve sensitivity of the instrument which will help improve their quality of life.

**Acknowledgements** Authors wish to acknowledge Dalal Al-Sowailam, Fatma Ridha, Haifa Al-Otaibi, Emad Hussain, Faisal Al-Munefi, Mohammad Abdulwahed, Mohamed Al-Fadhalah, and Talal Al-Meshri for their excellent work on data collection. Also, the authors wish to express deep gratitude to all patients for their participation in the study. Special thanks for Kuwait ministry of health for granting access to dialysis centers across the country.

**Funding** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Declarations

**Competing interests** Authors declare that they have no competing interests.

## References

- Kotronoulas GC, Papadopoulou CN, Papapetrou A, et al. Psychometric evaluation and feasibility of the Greek Pittsburgh Sleep Quality Index (GR-PSQI) in patients with cancer receiving chemotherapy. *Support Care Cancer*. 2011;19(11):1831–40.
- Akman T, Yavuzsen T, Sevgen Z, et al. Evaluation of sleep disorders in cancer patients based on Pittsburgh Sleep Quality Index. *Eur J Cancer Care (Engl)*. 2015;24(4):553–9.
- Nishiura M, Tamura A, Nagai H, et al. Assessment of sleep disturbance in lung cancer patients: relationship between sleep disturbance and pain, fatigue, quality of life, and psychological distress. *Palliat Support Care*. 2015;13(3):575–81.
- Zuurbier LA, Luik AI, Leening MJ, et al. Associations of heart failure with sleep quality: the Rotterdam Study. *J Clin Sleep Med*. 2015;11(2):117–21.
- Doi Y, Minowa M, Uchiyama M, et al. Psychometric assessment of subjective sleep quality using the Japanese version of the Pittsburgh Sleep Quality Index (PSQI-J) in psychiatric disordered and control subjects. *Psychiatry Res*. 2000;97(2–3):165–72.
- Ohayon MM, Roth T. Place of chronic insomnia in the course of depressive and anxiety disorders. *J Psychiatr Res*. 2003;37(1):9–15.
- Okamura H, Tsuda A, Yajima J, et al. Short sleeping time and psychobiological responses to acute stress. *Int J Psychophysiol*. 2010;78(3):209–14.
- Iseki K, Tsuruya K, Kanda E, et al. Effects of sleepiness on survival in Japanese hemodialysis patients: J-DOPPS study. *Nephron Clin Pract*. 2014;128(3–4):333–40.
- Han B, Zhu FX, Shi C, et al. Association between serum vitamin D levels and sleep disturbance in hemodialysis patients. *Nutrients*. 2017;9(2):139.
- Turkmen K, Erdur FM, Guney I, et al. Sleep quality, depression, and quality of life in elderly hemodialysis patients. *Int J Nephrol Renovasc Dis*. 2012;5:135–42.
- Brekke FB, Waldum B, Amro A, et al. Self-perceived quality of sleep and mortality in Norwegian dialysis patients. *Hemodial Int*. 2014;18(1):87–94.
- Unruh M, Kurella Tamura M, Larive B, et al. Impact of sleep quality on cardiovascular outcomes in hemodialysis patients: results from the frequent hemodialysis network study. *Am J Nephrol*. 2011;33(5):398–406.
- Hinz A, Glaesmer H, Brähler E, et al. Sleep quality in the general population: psychometric properties of the Pittsburgh Sleep Quality Index, derived from a German community sample of 9284 people. *Sleep Med*. 2017;30:57–63.
- Buysse DJ, Reynolds CF 3rd, Monk TH, et al. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193–213.
- Salahuddin M, Maru TT, Kumalo A, et al. Validation of the Pittsburgh sleep quality index in community dwelling Ethiopian adults. *Health Qual Life Outcomes*. 2017;15(1):58.
- Zhong QY, Gelaye B, Sánchez SE, et al. Psychometric Properties of the Pittsburgh Sleep Quality Index (PSQI) in a Cohort of Peruvian Pregnant Women. *J Clin Sleep Med*. 2015;11(8):869–77.
- Curcio G, Tempista D, Scarlata S, et al. Validity of the Italian version of the Pittsburgh Sleep Quality Index (PSQI). *Neurol Sci*. 2013;34(4):511–9.
- Suleiman KH, Yates BC, Berger AM, et al. Translating the Pittsburgh Sleep Quality Index into Arabic. *West J Nurs Res*. 2010;32(2):250–68.
- Farrahi J, Nakhrae N, Sheibani V, et al. Psychometric properties of the Persian version of the Pittsburgh Sleep Quality Index addendum for PTSD (PSQI-A). *Sleep Breath*. 2009;13(3):259–62.
- Burkhalter H, Sereika SM, Engberg S, et al. Structure validity of the Pittsburgh Sleep Quality Index in renal transplant recipients: A confirmatory factor analysis. *Sleep Biol Rhythms*. 2010;8:274–81.
- Hita-Contreras F, Martínez-López E, Latorre-Román PA, et al. Reliability and validity of the Spanish version of the Pittsburgh Sleep Quality Index (PSQI) in patients with fibromyalgia. *Rheumatol Int*. 2014;34(7):929–36.

22. Del Rio João KA, Becker NB, de Neves JS, et al. Validation of the Portuguese version of the Pittsburgh Sleep Quality Index (PSQI-PT). *Psychiatry Res.* 2017;247:225–9.
23. Sohn SI, Kim DH, Lee MY, et al. The reliability and validity of the Korean version of the Pittsburgh Sleep Quality Index. *Sleep Breath.* 2012;16(3):803–12.
24. Al Maqbal M, Dunwoody L, Rankin J, et al. Psychometric properties and cultural adaptation of sleep disturbance measures in Arabic-speaking populations: A systematic review. *J Sleep Res.* 2020;29(1): e12877.
25. Mollayeva T, Thurairajah P, Burton K, et al. The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. *Sleep Med Rev.* 2016;25:52–73.
26. R\_Core\_Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.
27. Ebadi A, Ausili D, Albatineh AN, et al. Psychometric evaluation of the farsi version of the self-care of diabetes inventory in Iranian patients with diabetes. *Diabetes Metab Syndr Obes.* 2019;12:2775–84.
28. Lim CR, Harris K, Dawson J, et al. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open.* 2015;5(7): e007765.
29. Ho R. Handbook of univariate and multivariate data analysis and interpretation with SPSS. Chapman & Hall/ CRC Press; 2006.
30. Cronbach LJ. Coefficient alpha and the internal structure of tests. *psychometrika.* 1951;16(3):297–334.
31. Dunn TJ, Baguley T, Brunnsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *Br J Psychol.* 2014;105(3):399–412.
32. Kline P. The handbook of psychological testing. Psychology Press; 2000.
33. Zinbarg RE, Revelle W, Yovel I, et al. Cronbach's Alpha, Revelle's Beta, McDonald's Omega: Their relations with each and two alternative conceptualizations of reliability. *Psychometrika.* 2005;70(1):123–33.
34. Eisinga R, Grotenhuis M, Pelzer B. The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *Int J Public Health.* 2013;58(4):637–42.
35. Frey BB. The SAGE encyclopedia of educational research, measurement, and evaluation. 2018. (Sage Publications).
36. Karimova G, Martin C. A psychometric evaluation of the Hospital Anxiety and Depression Scale during pregnancy. *Psychol Health Med.* 2003;8(1):89–103.
37. Rosseel Y. lavaan: An R package for structural equation modeling. *J Stat Softw.* 2012;48(2):1–36.
38. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull.* 1990;107(2):238–46.
39. Akaike H. Factor analysis and AIC. New York, NY: Springer; 1987.
40. Steiger JH. Structural model evaluation and modification: an interval estimation approach. *Multivariate Behav Res.* 1990;25(2):173–80.
41. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling.* 1999;6(1):1–55.
42. Browne MW, Cudeck R. Alternative ways of assessing model fit. *Sociol Methods Res.* 1992;21(2):230–58.
43. Byrne BM. Structural equation modeling with AMOS: basic concepts, applications, and programming (multivariate applications series). Edition n, editor. New York: Taylor & Francis Group; 2010.
44. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol Bull.* 1980;88(3):588.
45. Ullman J. Using Multivariate Statistic. Massachusetts: Allyn and Bacon: Needham Heights; 2001.
46. Hooper D, Coughlan J, Mullen M. Structural equation modelling: guidelines for determining model fit. *Electron J Bus Res Methods.* 6(1):53–60.
47. Tabachnick BG, Fidell LS, Ullman JB. Using multivariate statistics. Boston: Pearson; 2007.
48. Martin CR, Tweed AE, Metcalfe MS. A psychometric evaluation of the hospital anxiety and depression scale in patients diagnosed with end-stage renal disease. *Br J Clin Psychol.* 2004;43(Pt 1):51–64.
49. Dietch JR, Taylor DJ, Sethi K, et al. Psychometric evaluation of the PSQI in U.S. college students. *J Clin Sleep Med.* 2016;12(8):1121–9.
50. Barbosa Neto JB, Germain A, Mattos PF, et al. Psychometric properties of the Brazilian version of the Pittsburgh Sleep Quality Index Addendum for PTSD (PSQI-A). *Braz J Psychiatry.* 2014;36(4):330–5.
51. Rener-Sitar K, John MT, Bandyopadhyay D, et al. Exploration of dimensionality and psychometric properties of the Pittsburgh Sleep Quality Index in cases with temporomandibular disorders. *Health Qual Life Outcomes.* 2014;21(12):10.
52. Cole JC, Motivala SJ, Buysse DJ, et al. Validation of a 3-factor scoring model for the Pittsburgh sleep quality index in older adults. *Sleep.* 2006;29(1):112–6.
53. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;27(2):53–5.
54. Aloba OO, Adewuya AO, Ola BA, et al. Validity of the Pittsburgh Sleep Quality Index (PSQI) among Nigerian university students. *Sleep Med.* 2007;8(3):266–70.
55. Krijnen W. Algorithms for unweighted least-squares factor analysis. *Comput Stat Data Anal.* 1996;21(2):133–47.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.