SCIENTIFIC INVESTIGATIONS

Reliability and validity of the Pittsburgh Sleep Quality Index among frontline COVID-19 health care workers using classical test theory and item response theory

Ling Wang, MD[1,2]; Yong-Xi Wu, MD[1,3]; Yi-Qi Lin, MD[1,3]; Lin Wang, MD[1,4]; Zhao-Nan Zeng, MD[1,5]; Xiao-Liang Xie, MD[1,6]; Qiu-Yang Chen, MD[1,6]; Shi-Chao Wei, MD[1,3]

[1]Shengli Clinical Medical College of Fujian Medical University, Fuzhou, China; [2]Department of Pharmacy, Fujian Provincial Hospital, Fuzhou, China; [3]Department of Sleep Center, Fujian Provincial Hospital, Fuzhou, China; [4]Administrative Office, Fujian Provincial Hospital, Fuzhou, China; [5]Center for Experimental Research in Clinical Medicine, Fujian Provincial Hospital, Fuzhou, China; [6]The First Operating Room, Fujian Provincial Hospital, Fuzhou, China

Study Objectives: The applicability of sleep-related scales to frontline medical staff for the COVID-19 pandemic has not been fully proved, so sleep survey results lack credibility and accuracy, creating difficulties for the guidance and treatment of frontline medical staff with sleep disorders, which is not conducive to the prevention and control of COVID-19. This study sought to analyze the reliability and validity of the Pittsburgh Sleep Quality Index (PSQI) among frontline medical staff fighting the COVID-19 pandemic.

Methods: A network questionnaire survey was used to investigate the PSQI among frontline medical staff who fought COVID-19 in Wuhan, China from March 19 to April 15, 2020. Combined with classical test theory and item response theory, the content validity, internal consistency, construct validity, and other aspects of the PSQI were evaluated.

Results: According to classical test theory, content validity, criterion validity, and construct validity of the PSQI were good. But the internal consistency was better after the deletion of the "daytime dysfunction" subscale. With regard to item response theory, difficulty, the differential item function, and the Wright map performed well.

Conclusions: The original PSQI showed acceptable applicability in frontline COVID-19 medical staff, and its characteristics moderately improved after the "daytime dysfunction" subscale was removed.

Keywords: validity, Pittsburgh Sleep Quality Index, frontline health care workers, COVID-19

Citation: Wang L, Wu Y-X, Lin Y-Q, et al. Reliability and validity of the Pittsburgh Sleep Quality Index among frontline COVID-19 health care workers using classical test theory and item response theory. *J Clin Sleep Med*. 2022;18(2):541–551.

BRIEF SUMMARY

Current Knowledge/Study Rationale: The Pittsburgh Sleep Quality Index (PSQI) is one of the most frequently used scales in sleep quality investigation. However, the PSQI requires validation in medical staff, especially first-line medical staff against COVID-19. Here we explore the reliability and validity of PSQI among frontline COVID-19 health care workers using classical test theory and item response theory.

Study Impact: This is the first study that validates the PSQI in COVID-19 frontline health care workers, and shows that the PSQI is acceptable for application in this population. However, the exclusion of "daytime dysfunction" could improve the validity and reliability of the PSQI. The applicability of the PSQI still calls for further research, which should involve multiple centers and objective indicators such as sleep electroencephalograms.

INTRODUCTION

In the novel coronavirus outbreak, frontline health care is challenging, arduous, and demanding work that requires a high level of motivation and dedication. Many frontline health care workers for COVID-19 experience sleep loss as they adjust and cope with workloads and harsh environments. A meta-analysis found that 45.1% (95% confidence interval [CI], 37.2%–53.1%) of Chinese medical workers had low sleep quality during the pandemic, and the pooled mean total score of the Pittsburgh Sleep Quality Index (PSQI) was 9.83 (95% CI, 8.61–11.04). Health care workers in Wuhan, China had a higher total PSQI score than those in other regions (Cochrane Q statistic (Q) = 9.21, $P = .002$).[1] Several other Chinese studies found that compared

with non-frontline or nonmedical staff, the prevalence of sleep disorders in frontline COVID-19 medical staff is higher (16.4%–61.6%), with the epidemic of anxiety and depression.[2–6] Similarly, multiple investigations in Egypt, Turkey, Iran, and other Middle East regions[7–9]; Nigeria; Serbia; other African regions[10,11]; and Brazil[12] revealed that the proportion of poor sleep in medical staff fighting COVID-19 was greater than 45.7%, and the presence of anxiety features ($P < .001$) as well as depressive features ($P < .001$) could be used as significant independent predictors of poor sleep quality.[13] These studies suggest a high incidence of poor sleep in frontline COVID-19 health care workers. Unfortunately, poor sleep impairs memory and vigilance and, if chronic, may lead to related psychological illness and an increased risk of complications,[14,15] which may have a

dramatic impact on the day-to-day medical activity as well as overall health of frontline COVID-19 health care workers.

The PSQI, the most commonly used scale of sleep quality, was used in all the above studies. However, different factorial structures were found in various samples, casting doubt on the validity of the total PSQI score in detecting poor and good sleep quality.[16] Biases of the PSQI may be attributed to group differences (eg, age, disease, and culture), and further psychometric investigations in specific populations are needed.[17] Some characteristics of frontline COVID-19 medical staff, such as a strong work ethic, the large gap between knowledge of sleep and awareness of its importance, workload, and clinical needs, profoundly make their sleep pattern and schedule different, and perhaps there are other potential mechanisms that lead to their sleep disorders.[1–13,18] Therefore, validation of the PSQI in frontline COVID-19 health care workers is warranted and will support it as a clinical tool for this population.

However, there are few articles about the reliability and validity of the PSQI in medical personnel, or even in shift workers under high pressure. The reliability and validity of PSQI in regularly scheduled populations under low pressure, such as healthy people, community-dwelling adults, who were completely different from frontline COVID-19 health care workers, showed that Cronbach's α ranged from 0.69 to 0.84, but increased when a subscale (sleep medicine use) was excluded.[19–22] After 2–4 weeks, test-retest correlation coefficients for the 7 components ranged from 0.30 to 0.84.[19] The PSQI global score correlated with other sleep measures, such as the Insomnia Severity Index, Ford Insomnia Response to Stress Test, and Glasgow Sleep Effort Scale, but not with the Epworth Sleepiness Scale (ESS) or Snore, Tired, Observed, Pressure, Body mass index, Age, Neck, Gender (STOP-BANG).[23,24] The cutoff score was 5–8.5.[19–22] The PSQI provided a good and reliable distinction between normal and impaired sleep quality, but Becker et al[21] suggested excluding the "sleep medicine use" subscale and using it as a complementary qualitative assessment of health.

The reliability and validity research on PSQI is mainly based on the classical test theory (CTT). In most studies, Cronbach's α fluctuated between 0.70 and 0.85. No studies reported Cronbach's α above 0.9. With regard to the factor structure of the PSQI, most studies reported poorer fit in single-factor models and better fit in 2- or 3-factor models. The results were not consistent. The PSQI has a good correlation with the Insomnia Severity Index, polysomnography, and other sleep quality indicators, but weak or no correlation with conditions such as vomiting, anger, spasticity, and bladder dysfunction. The evidence on test-retest is limited and the appropriate length of time between test and retest is unclear, requiring further testing and definition. Gender differences lack consistent findings.[16,25,26] The advantage of CTT is that it reduces the impact of errors associated with each individual item by adding multiple items together. Although valuable, the use of CTT for examination and analysis of data is not without limits. First, all CTT-based statistics are sample dependent. CTT is only applicable if the participants' capabilities are comparable.[27,28] In addition, it is difficult for CTT to distinguish the cross-item common themes that are important to interest variables from the superficial

common themes.[27] Finally, CTT-based scales may tend to be more sensitive at the center (relative to the extreme) of the rating range. The scores at the tail are not estimated as accurately as the scores at the middle region.

Different from CTT, item response theory (IRT), with Rasch as the most commonly used model, is based on theories like item characteristic curve and latent trait theory. IRT makes up for the deficiencies of CTT in many aspects, such as the stability of item parameters and the estimation of measurement errors.[29] IRT describes the relationship between a latent trait and the probability of choosing a particular item response that indicates the trait being measured. IRT can enable researchers to create a visual representation of how well items are performing and therefore allows for an in-depth analysis at both the item level and the person level using the same metric.[30] There are 2 articles currently evaluating PSQI using IRT. Bellini et al[31] analyzed the individual items of PSQI in patients with irritable bowel syndrome by a unitary bowel-sleep model based on IRT. Revealing a strong positive correlation between the severity of irritable bowel syndrome symptoms and sleep disturbances. Chien et al[32] used Rasch analysis to validate the items of the revised PSQI (SC_PSQI) in Taiwan's high-tech workers. These results indicate that the SC_PSQI with a 0 to 2 scoring scale can be used as a one-dimensional scale to assess sleep quality. Although helpful in many cases, IRT also has clear limitations. Most applications of the IRT models cannot currently be conducted with commonly used statistical programs. Moreover, IRT will improve test validity by improving test reliability.[29] Therefore, CTT and IRT have complementary advantages. Both of these approaches can provide useful information at all stages of the evaluation.

The objective of the present study was to determine the validity and reliability of the PSQI in screening for sleep disorders among the frontline health care workers fighting against COVID-19. We analyzed content validity, criterion validity, internal consistency, and construct validity using CTT, as well as difficulty, differential item function, and Wright map using IRT in order to conduct a PSQI evaluation. This study provides a basis for improvement in and further research of the sleep quality of medical staff in the COVID-19 pandemic as well as other public health emergencies.

METHODS

Participants

Study inclusion criteria were as follows: (1) first-line qualified doctors or nurses in Wuhan for COVID-19, and (2) working time more than 1 week from March 19 to April 15, 2020. Exclusion features were as follows: (1) nervous system disease and psychological disorders, (2) women during their menstrual period, (3) thyroid dysfunction, (4) sleep apnea syndrome, and (5) dependence on sedative-hypnotics.

This study was performed in agreement with the Declaration of Helsinki ethics requirements and was approved by the Ethics Committee of the Faculty of Medicine, Fujian Provincial Hospital (K2020-03-145). Electronic signed informed consents were obtained from all participants.

Instruments

The frontline health care workers who fought against COVID-19 in Wuhan from March 19 to April 15, 2020, were invited to participate in a sleep survey via a smartphone-based WeChat (Tencent Corporation, Shenzhen, CN) application. Those who agreed to participate started the survey by clicking the enclosed website link or quick response (QR) code (2-dimensional bar code). Moreover, patients finished the survey without any assistance, while additional virtual guidance related to the administration of the survey, if asked for, was offered via WeChat platform or phone calls by individual members of our team. Participants were considered as enrolled if they returned the completed e-questionnaire. Obtained survey information including name, age, sex, score, and so on was collected by the Wen-Juan-Xing platform (Ranxing Information Technology Corporation, Changsha, CN), which is a free and open platform for survey design.

The content of the sleep survey contained general information (demographic and sociological data, psychological status, previous disease) and a Chinese version of sleep or health-related scales (PSQI, ESS, Morning and Evening Questionnaire [MEQ], and General Health Questionnaire [GHQ]).

The Chinese version of the PSQI consists of 19 self-rated questions grouped into 7 subscales: self-reported sleep quality, sleep latency, sleep duration, sleep efficiency, sleep disturbance, sleep medicine use, and daytime dysfunction. The score of each subscale was weighted equally on a 0–3 scale, and summed to yield a global PSQI score ranging from 0 to 21. Higher scores indicated worse sleep quality. The validity of the Chinese version of the PSQI has been supported in patients with cancer,[33] students,[34] and community-dwelling older adults.[35]

The ESS is a self-administered questionnaire, comprising 8 questions that survey the likelihood of daytime sleepiness. Each question is scored from 0 (would never doze) to 1 (slight chance of dozing), 2 (moderate chance of dozing), and 3 (high chance of dozing). The total score ranges from 0 to 24, with a score of 11–24 indicating excessive daytime sleepiness. The reliability and validity of the Chinese version of ESS have been investigated in Chinese patients with obstructive sleep apnea syndrome.[36,37]

The MEQ was used to classify participants into 3 different groups. Participants with scores lower than 42 were classified as evening types, participants with scores between 42 and 58 were classified as intermediate types, and participants with scores above 58 were classified as morning types.

The GHQ (GHQ-12) is a simple tool made for the identification of general psychiatric disorders. There are 12 questions about respondents' "depressive, anxiety symptoms, confidence, and overall happiness," measured on a Likert scale (0 = less than usual, 1 = no more than usual, 2 = rather more than usual, 3 = much more than usual). The scores for the 12 questions were summed, resulting in a value ranging from 0 (the least severe psychiatric disorder) to 36 (the most severe psychiatric disorder). A respondent with a GHQ-12 score > 18 was defined as having general psychiatric disorders. The Chinese version of the GHQ (GHQ-12 and GHQ-30) has been standardized for rural residents, parents of primary school children, young civil servants, and other general populations.[38–45]

CTT analysis

Content validity

Content validity demonstrates the coincidence degree between the actually measured content of a scale and the content to be measured. Six experts in the field of sleep were invited to assess the content relevance of the PSQI items by e-mail using a 4-point scoring system, with "no correlation" as 1 point, "weak correlation" as 2 points, "strong correlation" as 3 points, and "very strong correlation" as 4 points. The content validity index (CVI) was calculated on the item-level CVI (I-CVI), average scale-level CVI (S-CVI/Ave), and adjusted kappa value (K*). An I-CVI ≥ 0.78 and S-CVI ≥ 0.90 are considered good. The evaluation of K* was as follows: 0.40–0.59 for general, 0.60–0.74 for good, and > 0.74 for excellent.[46]

Criterion validity

Criterion validity is the extent to which a measure is related to an outcome. It is usually assessed through comparison with a well-established measurement that acts as the criterion against which the new instrument is assessed. The Pearson correlation analysis was performed on the total score and each subscale score of the PSQI with the total scores of ESS, MEQ-5, and GHQ-12, respectively. The receiver operating characteristic curve (ROC) of the PSQI for predicting sleep quality was drawn and the area under the ROC (AUC) was calculated based on self-reported sleep quality assessed by participants. The following benchmarks were used for AUC interpretation: 0.50 to 0.69 = "poor," 0.70 to 0.79 = "acceptable," 0.80 to 0.89 = "excellent," and ≥ 0.90 = "outstanding."[47] The optimal cutoff was identified as the PSQI score maximizing the Youden's J statistic (sensitivity + specificity − 1).

Construct validity

Construct validity evaluates the degree to which an instrument/test measures the intended construct. Exploratory factor analysis[48] and confirmatory factor analysis[49,50] were used.

Internal consistency

Internal consistency refers to the degree of consistency among all internal items. Cronbach's α coefficient was used to evaluate the reliability of internal consistency, and it is generally considered that a coefficient above 0.7 is good.[49]

Split-half reliability

PSQI subscales were divided into 2 equal groups (2 halves) by the even-odd grouping method. The correlation between the 2 halves was calculated by the Spearman Brown's formula. A correlation coefficient greater than 0.7 was generally considered good.[51–53]

Test-retest reliability

Test-retest reliability reflects the stability and consistency of the test across time. Using the same test method, some participants were tested twice, in turn, and the correlation coefficient of the scores obtained from the 2 tests was calculated. The intraclass correlation coefficient was usually used as the evaluation index of test-retest reliability. An intraclass correlation coefficient

of greater than 0.70 is considered indicative of acceptable reliability.[54] Two weeks later, 30 participants completed the questionnaire again, and the test-retest reliability was calculated.

IRT analysis

The Rasch model was used to analyze the reliability and validity of the PSQI. The unidimensionality of latent traits is an important prerequisite for Rasch model analysis. The principal component analysis of standardized residuals was used for a dimension test of the PSQI, and the eigenvalues between 1.4 and 2.1 were indicators of the unidimensionality of the scale.[55] The outfit mean square (outfit MNSQ) and infit MNSQ were used as the model fitting index, and the value range was 0.50–1.50. The closer to 1, the better the fit.[56] The difficulty level of each subscale was determined by average logit measure values (measure). The higher the measure value, the more difficult the subscale will be.[57,58] The point-measure correlation (PT-measure) indicated the correlation between individual performance on a subscale and that on the whole scale, and the range was 0.40–0.80.[59] Differential item function (DIF) was used to detect the differences of the same subscale in different participants' traits. DIF contrast greater than 0.5 logits and the absolute value of a T value greater than 2 indicated DIF.[56] The Wright map reflected the overall quality of the scale and demonstrated the joint distribution of a person's ability and item difficulty on the same line.[60] The reliability of Rasch was measured by the item reliability index and the individual reliability index. The closer the 2 reliability indexes were to 1, the better.[61]

Statistical analysis

SPSS 23.0 software (IBM Corporation, Armonk, NY) AMOS 22.0 software (IBM Corporation, Armonk, NY) were used to conduct CTT. Winsteps software (version 4.0.1 Winsteps.com, Beaverton, OR, USA) was used to perform the Rasch analysis.

RESULTS

Basic information

A total of 528 valid questionnaires were collected in this study. The participants included frontline COVID-19 doctors and nurses in Wuhan. The mean ± SD age was (33.63 ± 6.4) years old, including 139 males (26.3%) and 389 females (73.7%). The mean ± SD total PSQI score was 6.91 ± 3.8. Taking the total PSQI score > 7 as the screening criterion of sleep disorders,[62,63] there were 199 patients with sleep disorders, and the positive rate was 37.7%. Sociodemographic characteristics of the participants are shown in **Table 1**.

Reliability and validity of CTT

Content validity

The *S*-CVI/Ave of PSQI was 0.905, which was greater than 0.90, indicating that the overall content validity of the PSQI was good. Except for the score of "daytime dysfunction" subscale, the I-CVI of each subscale was greater than 0.78 and the K* was above 0.74 (**Table 2**), which showed that these

Table 1—Sociodemographic characteristics of the study cohort.

Variables	n (%)	PSQI Score	P Value
Age (y)			
<30	139 (26.3)	7.12 ± 3.5	.41
≥30	389 (73.7)	6.83 ± 3.9	
Sex			
Male	133 (25.2)	6.60 ± 3.8	.28
Female	395 (74.8)	7.01 ± 3.8	
Profession			
Doctor	114 (21.6)	6.32 ± 3.9	.06
Nurse	414 (78.4)	7.07 ± 3.7	

PSQI = Pittsburgh Sleep Quality Index.

subscales have excellent content validity, correlation, and representativeness.

Criterion validity

The ESS, MEQ-5, and GHQ-12 were used as criteria to evaluate the criterion validity of the PSQI. Except for sleep efficiency, the total score and each subscale score of the PSQI were positively correlated with the scores on the ESS and GHQ-12, and negatively correlated with the MEQ-5 score (*P* < .05; see **Table 3**). The "sleep efficiency" subscale score was positively correlated with ESS and GHQ-12 scores, but there was no statistical difference. The sleep efficiency score had a significant negative correlation with MEQ-5 scores (*P* < .05). This suggested that the PSQI has good criterion validity. Taking the self-rated sleep quality of survey participants as the standard, the ROC of the PSQI for predicting sleep quality was drawn. The AUC was 0.919 (95% confidence interval, 0.896–0.942). When the total score of PSQI was 6.5, the Youden index was the highest, indicating that participants with PSQI scores above this critical value had poorer sleep quality (**Figure 1**).

Construct validity

The Kaiser-Meyer-Olkin (KMO) value of 0.757 and Bartlett's sphericity test results ($\chi^2 = 950.189$, *P* < .001) showed that the

Table 2—Content validity of the PSQI.

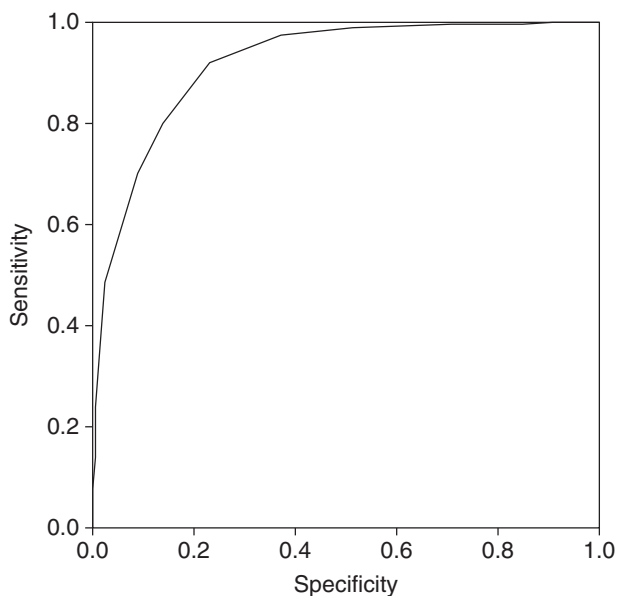| PSQI Subscales	I-CVI	Pc	K*
Self-reported sleep quality	0.833	0.09375	0.816
Sleep latency	1	0.01563	1
Sleep duration	1	0.01563	1
Sleep efficiency	1	0.01563	1
Sleep disturbance	1	0.01563	1
Sleep medicine use	0.833	0.09375	0.816
Daytime dysfunction	0.667	0.234375	0.565

I-CVI = item-level content validity index, K* = adjusted kappa value, Pc = the probability of chance universal agreement, PSQI = Pittsburgh Sleep Quality Index.

Table 3—Criterion validity of the PSQI.

PSQI Subscales	ESS	MEQ-5	GHQ-12
Self-reported sleep quality	0.256***	−0.298***	0.351***
Sleep latency	0.116**	−0.278***	0.246***
Sleep duration	0.157***	−0.101*	0.181***
Sleep efficiency	0.061	−0.159***	0.048
Sleep disturbance	0.395***	−0.220***	0.347***
Sleep medicine use	0.141**	−0.182***	0.199***
Daytime dysfunction	0.299***	−0.161***	0.411***
Total scores	0.302***	–0.309***	0.388***

ESS = Epworth Sleepiness Scale, GHQ-12 = 12-item General Health Questionnaire, MEQ-5 = 5-item Morning and Evening Questionnaire, PSQI = Pittsburgh Sleep Quality Index. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$.

Table 4—Exploratory factor analysis of the PSQI.

PSQI Subscales	Factor 1	Factor 2	Communalities
Self-reported sleep quality	0.711	0.426	0.688
Sleep latency	0.591	0.413	0.520
Sleep duration	0.179	0.772	0.629
Sleep efficiency	0.065	0.818	0.673
Sleep disturbance	0.676	0.236	0.513
Sleep medicine use	0.670	0.153	0.472
Daytime dysfunction	0.756	–0.226	0.622

PSQI = Pittsburgh Sleep Quality Index.

PSQI was suitable for exploratory factor analysis. In exploratory factor analysis (**Table 4**), 2 common factors were extracted, and the cumulative variance contribution rate was 58.828%. The communality of each subscale was greater than 0.4, revealing that subscales were well expressed by common factors. With regard to confirmatory factor analysis, it can be seen from **Table 5** that the 2-factor model had a good fit. Goodness of fit index (GFI), adjusted goodness of fit index (AGFI), comparative fit index (CFI), normed fit index (NFI) were all greater than 0.90, and standardized root-mean-square residual (SRMR) was less than 0.08. However, χ^2/df was more than 3, and root-mean-square error of approximation (RMSEA) was above 0. 08, indicating that the model still needed to be modified. The standardized estimate of the "daytime dysfunction" subscale in the 2-factor model was much smaller than that of other subscales, so the model was revised after the deletion of the "daytime dysfunction" subscale (**Figure 2**). All the indicators showed that the revised 2-factor model was better.

Internal consistency

The overall Cronbach's α coefficient of the PSQI was 0.746, indicating good internal consistency of the scale as a whole. The correlation coefficient between the score of each subscale and the adjusted PSQI total score (excluding the score of the corresponding subscale) was calculated to investigate the consistency between 1 subscale and others. It was found that the correlation coefficient of the "daytime dysfunction" subscale was too small, and the deletion of the "daytime dysfunction" subscale could increase the value of Cronbach's α (**Table 6**), indicating that the deletion of this subscale could improve the internal consistency of the PSQI.

Split-half reliability and test-retest reliability

The split-half reliability of the PSQI was good, with a correlation coefficient of 0.888. The intraclass correlation coefficient was within the range of 0.709–0.813, indicating that the test-retest reliability of the PSQI was also good (**Table 6**).[54]

Reliability and validity of IRT

Unidimensional analysis

The principal component analysis of standardized residuals showed that the variance attributable to the Rasch factor was good, with eigenvalues of 1.6, confirming the factors' unidimensionality. The data were suitable for analysis by the Rasch model.

Fitness and difficulty

The statistical results of the Rasch analysis are shown in **Table 7**. The outfit MNSQ and infit MNSQ of the PSQI were 0.65–1.44 and 0.68–1.27, respectively, showing a good fit. The PSQI could effectively and accurately depict the characteristics of participants. With regard to the difficulty level, the average logit measure value (measure) was from −1.59 to 1.55,

Figure 1—ROC of the PSQI to predict sleep quality.

PSQI = Pittsburgh Sleep Quality Index, ROC = receiver operating characteristic.

Table 5—Confirmatory factor analysis of the PSQI.

Model	Absolute Index					Relative Index		Parsimony Index
	χ^2/df	GFI	AGFI	SRMR	RMSEA	CFI	NFI	AIC
1-Factor	11.269	0.930	0.860	0.060	0.137	0.846	0.835	185.772
2-Factor	7.122	0.958	0.910	0.041	0.105	0.915	0.903	122.592
Revised 2-factor	2.965	0.984	0.921	0.019	0.066	0.976	0.966	103.3

AGFI = Adjusted Goodness of Fit Index, AIC = Akaike Information Criteria, CFI = Comparative Fit Index, GFI = Goodness of Fit Index, NFI = Normed Fit Index, PSQI = Pittsburgh Sleep Quality Index, RMSEA = Root-Mean-Square Error of Approximation, SRMR = Standardized Root-Mean-Square Residual.

showing moderate difficulty. The PT-measure ranged from 0.45 to 0.75, indicating a good correlation between individual performance in a certain subscale and that in the whole scale.

DIF

The DIF function was determined by measuring the stability for the completion of the PSQI between males and females, so as to demonstrate the universality of the PSQI. As shown in **Table 8**,

Figure 2—Factor loadings for 2-factor model without the subscale of daytime dysfunction.

T values were less than 2 for all subscales except for the "daytime dysfunction" subscale. The "daytime dysfunction" subscale had the largest T value (T = 2.20), but it did not reach a significant level. Therefore, there was no DIF contrast between sexes.
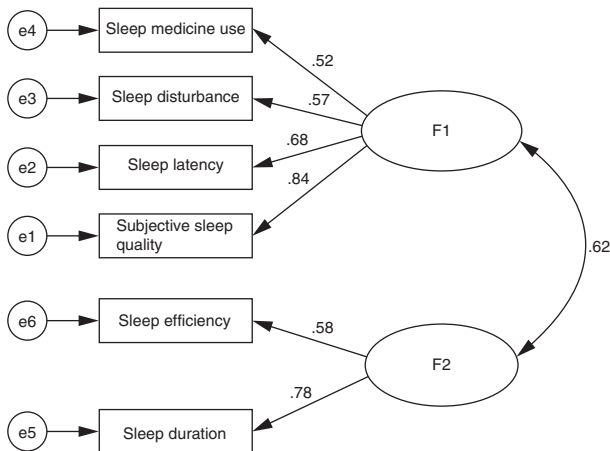
Adaptability between participant ability and item difficulty

The appropriateness of PSQI difficulty and participants' ability is shown in the Wright map (**Figure 3**). Participants are plotted on the left side of the map and subscales are on the right side. Higher-functioning participants (frontline health care workers with high ability to complete the scale) and more-difficult-to-perform subscales are at the top of the map. It can be seen that the participants' ability matched with the difficulty of subscales, mainly distributed in the lower position of the Wright map. The abilities of the subjects to complete PSQI were mainly plotted in the middle of the map, and there were few subjects with extremely high or extremely low abilities. Similarly, there was no extremely difficult subscale.

Reliability of the whole scale

The item reliability index was 0.99 and the person reliability index was 0.70. All subscales of the PSQI were consistent, and the survey results for different participants were stable.

DISCUSSION

High-quality data concerning the impact of the COVID-19 pandemic on the mental health of the entire population, especially

Table 6—The internal consistency and test-retest reliability of the PSQI.

PSQI Subscales	Cronbach's α if Subscale Deleted	Corrected Subscale—Total Correlation	Test-Retest Reliability
Self-reported sleep quality	0.675	0.685	0.731
Sleep latency	0.700	0.527	0.753
Sleep duration	0.723	0.431	0.802
Sleep efficiency	0.739	0.372	0.766
Sleep disturbance	0.710	0.525	0.813
Sleep medicine use	0.712	0.478	0.795
Daytime dysfunction	0.750	0.305	0.709

PSQI = Pittsburgh Sleep Quality Index.

Table 7—Subscale statistics of Rasch analysis.

PSQI Subscales	Total Score	Count	Measure	Model SE	Infit		Outfit		PT-Measure	
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.
Self-reported sleep quality	732	552	−0.95	0.08	0.68	−5.8	0.65	−5.9	0.75	0.60
Sleep latency	945	552	−1.59	0.06	0.96	−0.8	0.96	−0.7	0.73	0.71
Sleep duration	462	552	0.25	0.06	1.11	1.8	1.37	4.0	0.55	0.61
Sleep efficiency	401	552	0.33	0.06	1.22	3.1	1.42	3.2	0.52	0.58
Sleep disturbance	664	552	−0.31	0.08	0.87	−2.2	0.84	−2.5	0.67	0.59
Sleep medicine use	242	552	0.73	0.06	0.86	−1.7	0.71	−1.2	0.50	0.48
Daytime dysfunction	299	552	1.55	0.06	1.27	2.9	1.44	2.4	0.45	0.52

CORR = Correlation Coefficient, Exp = Expected Value, MNSQ = mean square, PSQI = Pittsburgh Sleep Quality Index, PT-measure = Point-Measure Correlation, SE = standard error, ZSTD = Probability of A Mean-Square Statistic Expressed as A Z-statistic.

health care professionals, are called for.[64] Insufficient sleep and poor sleep quality are direct risk factors for the mental and physical health of medical staff, and also potential key issues affecting medical risks.[65,66] Therefore, it is of great importance to explore the sleep status and its influencing factors on medical staff in public health emergencies, so as to carry out early intervention accordingly. However, a good instrument is a prerequisite for a successful investigation. Whether commonly used sleep scales in a clinic can correctly reflect the sleep status of frontline medical staff should be determined. The PSQI is widely used to investigate the sleep quality of various populations. However, due to sample limitations, the PSQI does not cover the sleep differences in different populations. The applicability of the PSQI to some characteristic groups, such as medical staff, has not been fully confirmed. This study analyzed the reliability and validity of the PSQI among frontline health care workers fighting against COVID-19, so as to provide evidence for further investigation on sleep quality of medical staff in the COVID-19 pandemic as well as other public health emergencies.

The results above show that the total PSQI score of frontline medical staff in COVID-19 was 6.91 ± 3.8. Those with scores >7 accounted for 37.7%, which was less than the 100%

Table 8—DIF of the PSQI.

PSQI Subscales	T	DIF Contrast
Self-reported sleep quality	1.89	0.33
Sleep latency	1.98	0.29
Sleep duration	1.91	0.28
Sleep efficiency	0.58	0.08
Sleep disturbance	0.00	0.00
Sleep medicine use	1.17	0.18
Daytime dysfunction	2.20	0.30

DIF = differential item function, PSQI = Pittsburgh Sleep Quality Index, T = the statistical probability of the DIF size relative to its measurement error expressed as an approximate unit-normal deviate.

reported by Wu and Wei,[67] but higher than the 30% reported by Cheng et al.[68] Although the conclusions of different studies differ, the number of participants in our study is larger than those reported in other articles concerning sleep status of frontline clinical staff against COVID-19. The prevalence of sleep disorders among frontline medical staff is higher than that among non-frontline and nonmedical staff. However, no more than 50% of frontline COVID-19 medical personnel had a PSQI score > 7 points, suggesting that the sleep quality of frontline COVID-19 medical personnel might be better than expected.

According to the results of CTT, the internal consistency and repeatability of the PSQI was good in frontline COVID-19 medical staff, but the related coefficients were lower than that in patients with cancer and community-dwelling older adults.[33,35] After elimination of "daytime dysfunction," the Cronbach's α coefficient increased, suggesting that "daytime dysfunction" might need to be excluded or revised. This is consistent with the studies on validity of the PSQI in older adults.[69–71]
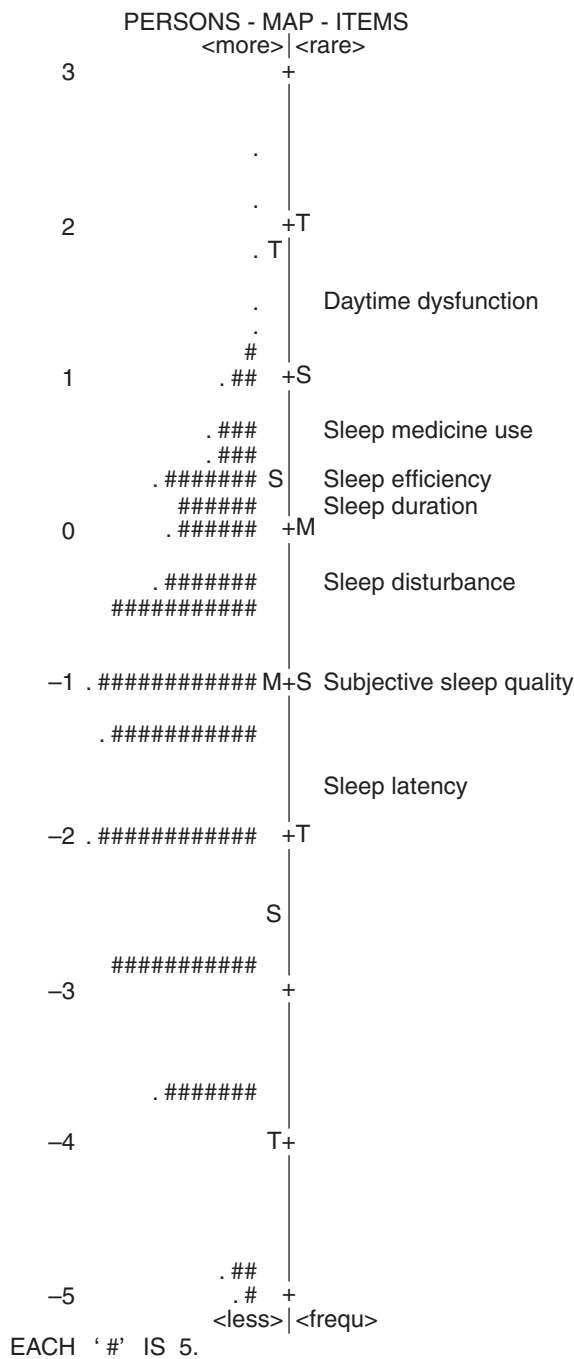
In terms of structural validity, the best model for the PSQI was the 3-factor model in older adults,[72] patients with chronic fatigue syndrome,[73] women with posttraumatic stress disorder,[74] and multiethnic Asian populations.[75] However, in breast cancer survivors,[76] the 2-factor model is superior. In this study, exploratory factor analysis and confirmatory factor analysis both supported the 2-factor model for frontline COVID-19 medical staff, but a 2-factor model with the deletion of "daytime dysfunction" was preferable. This was similar to improvements in content validity and internal consistency after the deletion of "daytime dysfunction." This was probably because the expression or scope of "daytime dysfunction" is unclear and it is easy to be answered incorrectly. On the other hand, daytime function is affected by complex factors not limited to sleep. Irregular shifts can affect sleep quality and individuals with high-level professional titles experience a higher incidence of sleep disorders than those with low-level titles. The higher the level of stress, the higher the PSQI scores on self-reported sleep quality, sleep disorder, and daytime dysfunction.[77] In addition, general negative emotions caused by the pandemic are associated with complaints of daytime

Figure 3—Rasch analysis person-item chart.


```
             PERSONS - MAP - ITEMS
               <more>|<rare>
     3              +
                    |
                    |
                .   |
                    |
                .   |
     2              +T
              .    T|
                .   |        Daytime dysfunction
                .   |
                #   |
     1         . ## +S
                    |
             . ###  |        Sleep medicine use
             . ###  |
         . ####### S|         Sleep efficiency
           ######   |        Sleep duration
     0      . ###### +M
                    |
            . ###### |        Sleep disturbance
        ###########|
    -1 . ########### M+S      Subjective sleep quality
                    |
         . ##########|
                    |
                    |        Sleep latency
    -2 . ########### +T
                    |
                   S|
        ###########|
    -3              +
                    |
         . #######  |
    -4             T+
                    |
                    |
         . ##       |
    -5       . #    +
               <less>|<frequ>
       EACH  ' # '  IS  5.
```

functioning and affective symptoms.[10,78,79] However, Rasskazova[80] pointed out that complaints of daytime functioning difficulties depended not so much on the level and content of anxiety but on the level of dysfunction that interferes with everyday activities during the period of COVID-19. Therefore, while applying the PSQI to the frontline COVID-19 medical population, the deletion or modification of the "daytime dysfunction" subscale may be considered.

With the ESS, MEQ-5, and GHQ-12 as criteria, the score on the "sleep efficiency" subscale had no statistical significance with ESS or GHQ-12 scores. Therefore, among the frontline COVID-19 medical personnel, sleep efficiency has poor correlation with drowsiness and mental health. The absence of a significant correlation between sleep efficiency and total ESS score has been reported in other literature, but no specific reasons or influencing factors were found.[81,82] For the lack of association between sleep efficiency and total GHQ-12 score, this may be due to several reasons. First, the GHQ-12 reflects physical health status in recent weeks and involves various issues such as sleep, stress, coping, mood, and self-evaluation. Second, a widely available, GHQ-12–based, single-item sleep measure was not an adequate substitute for a multi-item measure of overall sleep disturbance. Although the measures produced largely similar associations for key determinants of poor sleep, the discrepancies between responses must be considered when analyzing data from a measure that uses a single sleep item.[83] Last, only 2 underlying factors of mental health status in the GHQ-12 were associated with differences in sleep status.[84]

The ROC results showed that the PSQI with a critical value of 6.5 could well screen those with good or bad sleep quality among the frontline medical staff fighting the novel coronavirus. Since the PSQI score was an integer, the critical value was rounded to 7. Those with scores less than 7 have normal sleep, and those with scores of more than 7 experience sleep disorders. The higher the score is, the worse sleep quality is, which is consistent with previous reports.[62,63]

Similar to the conclusions of CTT, the results of Rasch analysis showed that the PSQI had good adaptability and reliability for frontline COVID-19 health care workers. However, our study and similar studies mainly adopted the self-reported indicators for sleep quality evaluation instead of objective indicators, such as sleep electroencephalogram. Thus, sleep scales with objective indexes remain to be developed.

One limitation of this study was that the data were collected during the pandemic and the participants' views may easily change due to various factors, such as remission or exacerbation of COVID-19. Another major limitation was that we were not able to have access in detail to objective measures such as polysomnography and the Multiple Sleep Latency Test for everyone in this population. Sleep disorder is, after all, a broad concept, and different individuals may have a different understanding of and answers to the self-report questions in the PSQI. The lack of details regarding diagnosis might induce homogeneity problems. In addition, we were not able to investigate whether the indicators and parameters related to the reliability and validity of PSQI were changeless within the group across sociodemographic factors, such as work place (intensive care unit, Fangcang shelter hospital,[85] emergency room, clinic, or others), or the adequacy of material supplies.

CONCLUSIONS

As a whole, the PSQI had good reliability and validity among frontline COVID-19 health care workers on the basis of CTT and IRT. The exclusion of "daytime dysfunction" could improve the content validity, construct validity, and internal

consistency of the PSQI. This study provides reference for the investigation of sleep among frontline COVID-19 medical staff and even in other health emergencies. Future studies may focus on the evaluation of the appropriateness of the PSQI excluding "daytime dysfunction" in order to provide a proper content validity assessment of the improved PSQI. In addition, future work should include multiregional studies with expanded sample sizes to further confirm the practical effect of the PSQI in more health care providers, such as dialysis centers, community health centers, nursing homes, substance abuse treatment providers, and rehabilitation providers.

ABBREVIATIONS

CTT, classical test theory
CVI, content validity index
DIF, Differential Item Function
ESS, Epworth Sleepiness Scale
GHQ-12, General Health Questionnaire
IRT, item response theory
MEQ, Morning and Evening Questionnaire
MNSQ, mean square
PSQI, Pittsburgh Sleep Quality Index
ROC, receiver operating characteristic

REFERENCES

1. Xia L, Chen C, Liu Z, et al. Prevalence of sleep disturbances and sleep quality in Chinese healthcare workers during the COVID-19 pandemic: a systematic review and meta-analysis. *Front Psychiatry.* 2021;12:646342.

2. Zhang J, Xu D, Xie B, et al. Poor-sleep is associated with slow recovery from lymphopenia and an increased need for ICU care in hospitalized patients with COVID-19: a retrospective cohort study. *Brain Behav Immun.* 2020;88:50–58.

3. Wang W, Song W, Xia Z, et al. Sleep disturbance and psychological profiles of medical staff and non-medical staff during the early outbreak of COVID-19 in Hubei Province, China. *Front Psychiatry.* 2020;11:733.

4. Wang S, Xie L, Xu Y, Yu S, Yao B, Xiang D. Sleep disturbances among medical workers during the outbreak of COVID-2019. *Occup Med (Lond).* 2020;70(5):364–369.

5. Tu ZH, He JW, Zhou N. Sleep quality and mood symptoms in conscripted frontline nurse in Wuhan, China during COVID-19 outbreak: a cross-sectional study. *Medicine (Baltimore).* 2020;99(26):e20769.

6. Qi J, Xu J, Li BZ, et al. The evaluation of sleep disturbances for Chinese frontline medical workers under the outbreak of COVID-19. *Sleep Med.* 2020;72:1–4.

7. Amra B, Salmasi M, Soltaninejad F, et al. Healthcare workers' sleep and mood disturbances during COVID-19 outbreak in an Iranian referral center. *Sleep Breath.* 2021;13:1–8.

8. Abbas A, Al-Otaibi T, Gheith OA, Nagib AM, Farid MM, Walaa M. Sleep quality among healthcare workers during the COVID-19 pandemic and its impact on medical errors: Kuwait experience. *Turk Thorac J.* 2021;22(2):142–148.

9. Saracoglu KT, Simsek T, Kahraman S, et al. The psychological impact of COVID-19 disease is more severe on intensive care unit healthcare providers: a cross-sectional study. *Clin Psychopharmacol Neurosci.* 2020;18(4):607–615.

10. Olagunju AT, Bioku AA, Olagunju TO, Sarimiye FO, Onwuameze OE, Halbreich U. Psychological distress and sleep problems in healthcare workers in a developing context during COVID-19 pandemic: implications for workplace wellbeing. *Prog Neuropsychopharmacol Biol Psychiatry.* 2021;110:110292.

11. Stojanov J, Malobabic M, Stanojevic G, Stevic M, Milosevic V, Stojanov A. Quality of sleep and health-related quality of life among health care professionals treating patients with coronavirus disease-19. *Int J Soc Psychiatry.* 2021;67(2):175–181.

12. Brito-Marques JMAM, Franco CMR, Brito-Marques PR, Martinez SCG, Prado GFD. Impact of COVID-19 pandemic on the sleep quality of medical professionals in Brazil. *Arq Neuropsiquiatr.* 2021;79(2):149–155.

13. Abdellah MM, Khalil MF, Alhafiz A. Prevalence of poor sleep quality among physicians during the COVID-19 pandemic. *Cureus.* 2021;13(1):e12948.

14. Palmer CA, Alfano CA. Sleep and emotion regulation: an organizing, integrative review. *Sleep Med Rev.* 2017;31:6–16.

15. Cappuccio FP, Miller MA. Sleep and cardio-metabolic disease. *Curr Cardiol Rep.* 2017;19(11):110.

16. Fabbri M, Beracci A, Martoni M, Meneo D, Tonetti L, Natale V. Measuring subjective sleep quality: a review. *Int J Environ Res Public Health.* 2021;18(3):1082.

17. Larche CL, Plante I, Roy M, Ingelmo PM, Ferland CE. The Pittsburgh Sleep Quality Index: reliability, factor structure, and related clinical factors among children, adolescents, and young adults with chronic pain. *Sleep Disord.* 2021;2021:5546484.

18. Waage S, Pallesen S, Vedaa Ø, et al. Sleep patterns among Norwegian nurses between the first and second wave of the COVID-19 pandemic. *BMC Nurs.* 2021;20(1):105.

19. Sohn SI, Kim DH, Lee MY, Cho YW. The reliability and validity of the Korean version of the Pittsburgh Sleep Quality Index. *Sleep Breath.* 2012;16(3):803–812.

20. Curcio G, Tempesta D, Scarlata S, et al. Validity of the Italian version of the Pittsburgh Sleep Quality Index (PSQI). *Neurol Sci.* 2013;34(4):511–519.

21. Becker NB, de Neves Jesus S. Adaptation of a 3-factor model for the Pittsburgh Sleep Quality Index in Portuguese older adults. *Psychiatry Res.* 2017;251:298–303.

22. Del Rio João KA, Becker NB, de Neves Jesus S, Isabel Santos Martins R. Validation of the Portuguese version of the Pittsburgh Sleep Quality Index (PSQI-PT). *Psychiatry Res.* 2017;247:225–229.

23. Gomes AA, Marques DR, Meiavia AM, et al. Psychometric properties and accuracy of the European Portuguese version of the Pittsburgh Sleep Quality Index in clinical and non-clinical samples. *Sleep Biol Rhythms.* 2018;16(4):413–422.

24. Kotronoulas GC, Papadopoulou CN, Papapetrou A, Patiraki E. 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):1831–1840.

25. Mollayeva T, Thurairajah P, Burton K, Mollayeva S, Shapiro CM, Colantonio A. 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. Manzar MD, BaHammam AS, Hameed UA, et al. Dimensionality of the Pittsburgh Sleep Quality Index: a systematic review. *Health Qual Life Outcomes.* 2018;16(1):89.

27. DeVellis RF. Classical test theory. *Med Care.* 2006;44(11 Suppl 3):S50–S59.

28. Kohli N, Koran J, Henn L. Relationships among classical test theory and item response theory frameworks via factor analytic models. *Educ Psychol Meas.* 2015;75(3):389–405.

29. Thomas ML. The value of item response theory in clinical assessment: a review. *Assessment.* 2011;18(3):291–307.

30. Stover AM, McLeod LD, Langer MM, Chen WH, Reeve BB. State of the psychometric methods: patient-reported outcome measure development and refinement using item response theory. *J Patient Rep Outcomes.* 2019;3(1):50.

31. Bellini M, Gemignani A, Gambaccini D, et al. Evaluation of latent links between irritable bowel syndrome and sleep quality. *World J Gastroenterol.* 2011;17(46):5089–5096.

32. Chien TW, Hsu SY, Tai C, Guo HR, Su SB. Using Rasch analysis to validate the revised PSQI to assess sleep disorders in Taiwan's hi-tech workers. *Community Ment Health J.* 2008;44(6):417–425.

33. Ho RT, Fong TC. Factor structure of the Chinese version of the Pittsburgh Sleep Quality Index in breast cancer patients. *Sleep Med.* 2014;15(5):565–569.

34. Zheng B, Li M, Wang KL, Lv J. Analysis of the reliability and validity of the Chinese version of Pittsburgh sleep quality index among medical college students. Article in Chinese. *Beijing Da Xue Xue Bao.* 2016;48(3):424–428.

35. Lin XL, Lu DL, Gottschling J, Segal DL, Tang SY. Validation of a Chinese version of the Geriatric Anxiety Scale among community-dwelling older adults in mainland China. *J Cross Cult Gerontol.* 2017;32(1):57–70.

36. Chung KF. Use of the Epworth Sleepiness Scale in Chinese patients with obstructive sleep apnea and normal hospital employees. *J Psychosom Res.* 2000; 49(5):367–372.

37. Chen NH, Johns MW, Li HY, et al. Validation of a Chinese version of the Epworth Sleepiness Scale. *Qual Life Res.* 2002;11(8):817–821.

38. Guan M, Han B. Factor structures of General Health Questionnaire-12 within the number of kins among the rural residents in China. *Front Psychol.* 2019;10: 1774.

39. Liang Y, Wang L, Yin X. The factor structure of the 12-item General Health Questionnaire (GHQ-12) in young Chinese civil servants. *Health Qual Life Outcomes.* 2016;14(1):136.

40. Sun J, Stewart D, Yuan BJ, Zhang SH. Validation and normalization of the General Health Questionnaire 30 in parents with primary school children in China. *Compr Psychiatry.* 2012;53(5):593–599.

41. Ip WY, Martin CR. Psychometric properties of the 12-item General Health Questionnaire (GHQ-12) in Chinese women during pregnancy and in the postnatal period. *Psychol Health Med.* 2006;11(1):60–69.

42. Shek DT, Tsang SK. Reliability and factor structure of the Chinese GHQ-30 for parents with preschool mentally handicapped children. *J Clin Psychol.* 1995;51(2): 227–234.

43. Shek DT. Reliability and factorial structure of the Chinese version of the General Health Questionnaire. *J Clin Psychol.* 1987;43(6):683–691.

44. Chan DW. The Chinese version of the General Health Questionnaire: does language make a difference? *Psychol Med.* 1985;15(1):147–155.

45. Chan DW, Chan TS. Reliability, validity and the structure of the General Health Questionnaire in a Chinese context. *Psychol Med.* 1983;13(2):363–371.

46. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health.* 2007;30(4):459–467.

47. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol.* 2010;5(9):1315–1316.

48. Tefera GM, Megersa WA, Gadisa DA. Health-related quality of life and its determinants among ambulatory patients with epilepsy at Ambo General Hospital, Ethiopia: using WHOQOL-BREF. *PLoS One.* 2020;15(1):e0227858.

49. Dunleavy G, Bajpai R, Tonon AC, et al. Examining the factor structure of the Pittsburgh Sleep Quality Index in a multi-ethnic working population in Singapore. *Int J Environ Res Public Health.* 2019;16(23):4590.

50. Zhao Q, Yang C, Tang S, et al. Developing and testing the reliability and validity of the Brief Haze Weather Health Protection Behavior Assessment Scale-Adolescent Version (BHWHPBAS-AV). *Front Pediatr.* 2020;8:498885.

51. Ballou S, Gray S, Palsson OS. Validation of the Pandemic Emotional Impact Scale. *Brain Behav Immun Health.* 2020;9:100161.

52. Wang Y, Li K, Li H, et al. Development, reliability, and validity of the Home Blood Pressure Monitoring Adherence Scale for patients with chronic kidney disease. *Patient Prefer Adherence.* 2020;14:1863–1872.

53. Al-Dwaikat TN, Ta'an W, Alrawashdeh M, Baker NA, Al Ali NM. Development and psychometric evaluation of nurses and midwives' perceptions of their roles in primary healthcare. *Int J Nurs Sci.* 2020;7(4):460–465.

54. Bartko JJ. The intraclass correlation coefficient as a measure of reliability. *Psychol Rep.* 1966;19(1):3–11.

55. Raîche G. Critical eigenvalue sizes (variances) in standardized residual principal components analysis (PCA) [updated 2005]. https://www.rasch.org/rmt/rmt191h. htm. Accessed October 5, 2021.

56. Linacre JM, Wright BD. Construction of measures from many-facet data. *J Appl Meas.* 2002;3(4):486–512.

57. Lestari WT, Saputro S, Masykuri M, et al. Item analysis of teachnological pedagogical content knowledge (TPACK) in pre-service chemistry teachers using the Rasch model application. Paper published in: *Journal of Physics:* Conference Series, Volume 1511, International Conference in Science Education and Technology; November 23, 2019; Surakarta, Indonesia. Accessed October 5, 2021.

58. Seamon BA, Kautz SA, Velozo CA. Rasch analysis of the Activities-Specific Balance Confidence Scale in individuals poststroke. *Arch Rehabil Res Clin Transl.* 2019;1(3-4):100028.

59. Yan Z, Bond TG. Developing a Tasch measurement physical fitness scale for Hong Kong primary school-aged students. *Meas Phys Educ Exerc Sci.* 2011; 15(3):182–203.

60. Chen TA, O'Connor TM, Hughes SO, et al. TV parenting practices: is the same scale appropriate for parents of children of different ages? *Int J Behav Nutr Phys Act.* 2013;10:41.

61. Moeini S, Rasmussen JV, Klausen TW, Brorson S. Rasch analysis of the Western Ontario Osteoarthritis of the Shoulder index—the Danish version. *Patient Relat Outcome Meas.* 2016;7:173–181.

62. Xiao M, Huang G, Feng L, et al. Impact of sleep quality on post-stroke anxiety in stroke patients. *Brain Behav.* 2020;10(12):e01716.

63. 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). *Br J Psychiatry.* 2014;36(4):330–335.

64. Holmes EA, O'Connor RC, Perry VH, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry.* 2020;7(6):547–560.

65. Hassani S, Rahnama N, Seyedmehdi SM, et al. Association between occupational accidents and sleep apnea in hospital staff. *Tanaffos.* 2015;14(3):201–207.

66. Pappa S, Ntella V, Giannakas T, Giannakoulis VG, Papoutsi E, Katsaounou P. Prevalence of depression, anxiety, and insomnia among healthcare workers during the COVID-19 pandemic: a systematic review and meta-analysis. *Brain Behav Immun.* 2020;88:901–907.

67. Wu K, Wei X. Analysis of psychological and sleep status and exercise rehabilitation of front-line clinical staff in the fight against COVID-19 in China. *Med Sci Monit Basic Res.* 2020;26:e924085.

68. Cheng FF, Zhan SH, Xie AW, et al. Anxiety in Chinese pediatric medical staff during the outbreak of coronavirus disease 2019: a cross-sectional study. *Transl Pediatr.* 2020;9(3):231–236.

69. Zhang C, Zhang H, Zhao M, et al. Reliability, validity, and factor structure of Pittsburgh Sleep Quality Index in community-based centenarians. *Front Psychiatry.* 2020;11:573530.

70. Spira AP, Beaudreau SA, Stone KL, et al; Osteoporotic Fractures in Men Study. Reliability and validity of the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Scale in older men. *J Gerontol A Biol Sci Med Sci.* 2012;67(4): 433–439.

71. Beaudreau SA, Spira AP, Stewart A, et al; Study of Osteoporotic Fractures. Validation of the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Scale in older black and white women. *Sleep Med.* 2012;13(1):36–42.

72. Cole JC, Motivala SJ, Buysse DJ, Oxman MN, Levin MJ, Irwin MR. Validation of a 3-factor scoring model for the Pittsburgh Sleep Quality Index in older adults. *Sleep.* 2006;29(1):112–116.

73. Mariman A, Vogelaers D, Hanoulle I, Delesie L, Tobback E, Pevernagie D. Validation of the three-factor model of the PSQI in a large sample of chronic fatigue syndrome (CFS) patients. *J Psychosom Res.* 2012;72(2):111–113.

74. Casement MD, Harrington KM, Miller MW, Resick PA. Associations between Pittsburgh Sleep Quality Index factors and health outcomes in women with posttraumatic stress disorder. *Sleep Med.* 2012;13(6):752–758.

75. Koh HW, Lim RB, Chia KS, Lim WY. The Pittsburgh Sleep Quality Index in a multi-ethnic Asian population contains a three-factor structure. *Sleep Breath.* 2015;19(4):1147–1154.

76. Otte JL, Rand KL, Carpenter JS, Russell KM, Champion VL. Factor analysis of the Pittsburgh Sleep Quality Index in breast cancer survivors. *J Pain Symptom Manage.* 2013;45(3):620–627.

77. Li X, Gao X, Liu J. Cross-sectional survey on the relationship between occupational stress, hormone levels, and the sleep quality of oilfield workers in Xinjiang, China. *Int J Environ Res Public Health.* 2019;16(18):3316.

78. Simor P, Polner B, Báthori N, et al. Home confinement during the COVID-19: day-to-day associations of sleep quality with rumination, psychotic-like experiences, and somatic symptoms. *Sleep.* 2021;44(7):zsab029.

79. Gupta R, Grover S, Basu A, et al. Changes in sleep pattern and sleep quality during COVID-19 lockdown. *Indian J Psychiatry.* 2020;62(4):370–378.

80. Rasskazova EI. [Psychological factors of sleep and daytime complaints during the COVID-19 lockdown: the role of anxiety, well-being, autonomy, and coping]. *Zh Nevrol Psikhiatr Im S S Korsakova.* 2021;121(4. Vyp. 2):24–30.

81. Veqar Z, Hussain ME. Psychometric analysis of Epworth Sleepiness Scale and its correlation with Pittsburgh Sleep Quality Index in poor sleepers among Indian university students. *Int J Adolesc Med Health.* 2018;31(2):20160151.

82. Farah NM, Saw Yee T, Mohd Rasdi HF. Self-reported sleep quality using the Malay version of the Pittsburgh Sleep Quality Index (PSQI-M) in Malaysian adults. *Int J Environ Res Public Health.* 2019;16(23):4750.

83. Lallukka T, Dregan A, Armstrong D. Comparison of a sleep item from the General Health Questionnaire-12 with the Jenkins Sleep Questionnaire as measures of sleep disturbance. *J Epidemiol.* 2011;21(6):474–480.

84. Suzuki H, Kaneita Y, Osaki Y, et al. Clarification of the factor structure of the 12-item General Health Questionnaire among Japanese adolescents and associated sleep status. *Psychiatry Res.* 2011;188(1):138–146.

85. Chen S, Zhang Z, Yang J, Wang J, Zhai X, Bärnighausen T, Wang C. Fangcang shelter hospitals: a novel concept for responding to public health emergencies. *Lancet.* 2020;395(10232):1305–1314.

SUBMISSION & CORRESPONDENCE INFORMATION

Address correspondence to: Shi-Chao Wei, MD, Department of Sleep Center, Fujian Provincial Hospital, Fuzhou, China; Email: wei0506sc@126.com

DISCLOSURE STATEMENT