JCSM
Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

Psychometric Evaluation of the PSQI in U.S. College Students

Jessica R. Dietch, MS[1]; Daniel J. Taylor, PhD, D,ABSM, CBSM[1]; Kevin Sethi, PhD[1]; Kimberly Kelly, PhD[1]; Adam D. Bramoweth, PhD[2]; Brandy M. Roane, PhD, CBSM[3]

[1]Department of Psychology, University of North Texas, Denton, TX; [2]VISN 4 Mental Illness Research, Education and Clinical Center, VA Pittsburgh Healthcare System, Pittsburgh, PA; [3]University of North Texas Health Science Center, Department of Internal Medicine, Fort Worth, TX

Study Objectives: Examine the psychometric properties of the PSQI in two U.S. college samples.
Methods: Study I assessed convergent and divergent validity in 866 undergraduates who completed a sleep diary, PSQI, and other sleep and psychosocial measures. Study II assessed PSQI insomnia diagnostic accuracy in a separate sample of 147 healthy undergraduates with and without insomnia.
Results: The PSQI global score had only moderate convergent validity with sleep diary sleep efficiency (prospective global measure of sleep continuity; $r = 0.53$), the Insomnia Severity Index ($r = 0.63$), and fatigue ($r = 0.44$). The PSQI global score demonstrated good divergent validity with measures of excessive daytime sleepiness ($r = 0.18$), circadian preference ($r = -0.08$), alcohol ($r = 0.08$) and marijuana ($r = 0.05$) abuse scales, and poor divergent validity with depression ($r = 0.48$), anxiety ($r = 0.40$), and perceived stress ($r = 0.33$). Examination of other analogous PSQI and sleep diary components showed low to moderate convergent validity: sleep latency ($r = 0.70$), wake after sleep onset ($r = 0.37$), sleep duration ($r = 0.51$), and sleep efficiency ($r = -0.32$). Diagnostic accuracy of the PSQI to detect insomnia was very high (area under the curve = 0.999). Sensitivity and specificity were maximized at a cutoff of 6.
Conclusions: The PSQI demonstrated moderate convergent validity compared to measures of insomnia and fatigue and good divergent validity with measures of daytime sleepiness, circadian phase preference, and alcohol and marijuana use. The PSQI demonstrated considerable overlap with depression, anxiety, and perceived stress. Therefore, caution should be used with interpretation.

Keywords: PSQI, sleep quality, students, college, psychometrics, validation, undergraduate, insomnia
Citation: Dietch JR, Taylor DJ, Sethi K, Kelly K, Bramoweth AD, Roane BM. Psychometric evaluation of the PSQI in U.S. college students. *J Clin Sleep Med* 2016;12(8):1121–1129.

INTRODUCTION

The Pittsburgh Sleep Quality Index (PSQI)[1] is a self-report questionnaire that is widely used by clinicians and researchers to broadly assess several dimensions of sleep.[1] The PSQI has been validated in numerous populations and languages[2–4] as well as many medical populations such as patients with insomnia, patients with traumatic brain injury, patients with cancer, patients receiving bone marrow and renal transplants, nursing home residents, and pregnant women.[5–11] Although previously validated in young adults[12] and international college populations,[13,14] to our knowledge no studies have specifically examined the psychometric properties of the PSQI in a U.S. college sample.

For many college students, the transition from high school poses many new challenges, such as roommates/dorm life, scheduling difficulties, active social lives, work obligations, and a new level of independence. This transition may provide the necessary stress for development or exacerbation of a sleep disorder, such as insomnia, which is highly prevalent in this population.[15] The consequences of poor sleep in this population are increasingly salient, as several recent studies have shown poor sleep, defined broadly, is associated with poor physical and psychological health outcomes such as mood disturbances and decreased cognitive performance.[16–19] Further, poor sleep has been associated with worse academic performance,[20] which can result in dropout or underemployment. Therefore, determining the validity of measures to accurately screen and identify sleep disturbance and sleep disorders in this population is a public health imperative.

Furthermore, the PSQI is a broad subjective measure that purports to assess several domains of sleep, such as sleepiness, sleep quality, insomnia symptoms, and sleep medication use, but few studies[14,21] have examined accuracy with which the PSQI measures these specific sleep constructs. A previous study in a middle-aged adult sample demonstrated the PSQI is

BRIEF SUMMARY
Current Knowledge/Study Rationale: The Pittsburgh Sleep Quality Index (PSQI) is a self-report questionnaire that is widely used by clinicians and researchers to broadly assess several dimensions of sleep, but it has never been validated in U.S. college students. The current study aimed to assess which aspects of sleep and psychosocial factors the PSQI can accurately measure in U.S. college students, examine the diagnostic accuracy of the PSQI for Insomnia Disorder, and provide recommendations about the uses and limitations of the PSQI in this population.
Study Impact: If the PSQI is to be used in a college population as an insomnia screening tool, a cutoff score ≥ 6 is recommended in order to maintain high sensitivity of the test and maximize the diagnostic accuracy. The PSQI may be a good screening tool, but is not a substitute for sleep disorder diagnoses established through a clinical interview performed by an experienced sleep clinician—and users of the PSQI in populations with a history of psychiatric illness should exercise caution in interpretation.

Table 1—Sample characteristics for study I and study II.

Study I Characteristic	n (%) or mean ± SD
Age	20.4 (4.1)
Sex	
Male	217 (25.1%)
Female	622 (71.8%)
Race/Ethnicity	
Black or African American	107 (12.4%)
White	582 (67.2%)
Asian	48 (5.5%)
Hispanic	86 (9.9%)
Other	41 (5.0%)
Sleep Diary	
Sleep Latency (minutes)	21.5 (21.0)
Wake after Sleep Onset (minutes)	7.1 (10.4)
Total Sleep Time (hours)	7.48 (1.1)
Sleep Efficiency (%)	90.8 (6.1)
Insomnia Severity Index	7.2 (5.1)
Epworth Sleepiness Scale	8.6 (3.7)
Multidimensional Fatigue Inventory – General Fatigue	11.7 (3.5)
Morningness-Eveningness Questionnaire	16.0 (4.4)
Quick Inventory of Depressive Symptoms	6.6 (3.9)
State-Trait Anxiety Inventory	39.4 (10.5)
Perceived Stress Scale	18.1 (7.3)
Alcohol Use Disorder Identification Test	4.2 (4.7)
Marijuana Problems Scale	1.0 (2.9)
Study II Characteristic	**n (%) or mean ± SD**
Age	20.2 (2.6)
Sex	
Male	60 (40.8%)
Female	87 (59.2%)
Race	
Black or African American	17 (11.6%)
White	96 (65.3%)
Asian	5 (3.4%)
Biracial or Other	29 (19.8%)
Ethnicity	
Hispanic/Latino	43 (29.3%)
Not Hispanic/Latino	102 (69.4%)
Insomnia Diagnosis	
Insomnia	68 (46.3%)
No Insomnia	79 (53.7%)
Pittsburgh Sleep Quality Index global score	6.34 (4.5)
Insomnia Group	10.65 (2.79)
No Insomnia Group	2.63 (1.29)

weakly correlated with a measure of sleepiness, and uncorrelated with objective measures of sleep.[21] A study of Taiwanese college students found higher PSQI scores in individuals reporting symptoms of insomnia and poor sleep quality than in those reporting daytime sleepiness.[14] The current study aimed to assess which aspects of sleep and psychosocial factors the PSQI can accurately measure in U.S. college students, examine the diagnostic accuracy of the PSQI for Insomnia Disorder, and provide recommendations about the uses and limitations of the PSQI in this population.

STUDY I

The aim of Study I was to explore the psychometric properties of the PSQI in a large undergraduate sample, particularly examining convergent validity with one week of sleep diaries and several well-validated self-reported measures of sleep and daytime functioning (insomnia and fatigue). Divergent validity was examined with several well-validated self-report measures of sleep and psychosocial factors (sleepiness, circadian preference, depression, anxiety, perceived stress, alcohol abuse, and marijuana abuse).

Methods

Sample

A total of 1,039 undergraduate students volunteered to complete study questionnaires in exchange for credit in psychology courses. One hundred seventy-three participants were excluded for missing sleep diary or PSQI data. The final sample size was 866 students with an average age of 20.4 years (standard deviation [SD] 4.1) and 74.1% were females (3% did not provide gender information). Sample characteristics are presented in **Table 1**. The racial and ethnic diversity of the sample was similar to the overall student population: 67.4% Non-Hispanic White, 12.4% African American/Black, 10% Hispanic, 5.6% Asian/Pacific Islander, and 4.8% Other. This study was approved by the university institutional review board, and informed consent was obtained for all participants.

Measurement

Pittsburgh Sleep Quality Index (PSQI): The PSQI version used in the current study was a 19-item self-report retrospective questionnaire of the past 7 days designed to measure 7 domains called component scores: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medications, and daytime dysfunction. Component scores range from 0 (no difficulty) to 3 (severe difficulty), and, when summed, produce a global score ranging from 0 to 21. Scores > 5 indicate significant sleep disturbance. In the original validation study, this cutoff score successfully identified 84% of patients with sleep initiation/maintenance disorders, 89% of patients with disorders of excessive daytime sleepiness, and 97% of patients with depression.[1] Good psychometric properties have been established.[1] The Cronbach α for the PSQI in the current sample has previously been reported as 0.83,[22] which is identical to the value reported in the original validation study.[1]

Sleep Diary: Participants were asked to complete a sleep diary each morning upon awakening for 7 days, after which the following questionnaires were completed. The sleep diary is a subjective measure that asks participants to give an estimate of their sleep on the previous night (e.g., bed time, sleep onset latency, wake time). These variables allow for the calculation of additional sleep metrics like sleep latency (SL), sleep efficiency (SE), total sleep time (TST), and wake after sleep onset (WASO). Overall, sleep diaries have shown to correlate moderately well with both polysomnography and actigraphy, two well-validated objective measures of sleep.[23,24] The Consensus Sleep Diary[25] was not used, because it had not been developed at the time this data was collected, but the version of the sleep diary used in the current study was similar in most ways to the Consensus Sleep Diary. For a more detailed description of the sleep diary used in the current study, please see other publications that used the same diary.[15,26,27]

INSOMNIA SEVERITY INDEX (ISI): The ISI[28] is a self-report measure designed to assess perceived severity of insomnia over the past two weeks. The measure consists of 7 items on a 5-point Likert scale ranging from 0 (not at all satisfied) to 4 (very much satisfied). Total scores range from 0 to 28, with higher scores representing greater perceived insomnia severity. In the current study, the ISI yielded a Cronbach α of 0.86.

MULTIDIMENSIONAL FATIGUE INVENTORY – GENERAL FATIGUE SCALE (MFI): The MFI[29] is a self-report measure designed to assess fatigue across several dimensions in the past week. Total score calculation is not recommended, so the General Fatigue subscale was used in the current study to represent overall fatigue. The General Fatigue subscale consists of 4 items on a 5-point Likert scale ranging from 1 (yes, that is true) to 5 (no, that is not true). Total scores range from 4 to 20, with higher scores indicating greater fatigue. In the current study, the General Fatigue subscale of the MFI yielded a Cronbach α of 0.76.

EPWORTH SLEEPINESS SCALE (ESS): The ESS[30] is a self-report measure designed to assess the level of daytime sleepiness in recent times. The measure consists of 8 items on a 4-point Likert scale on which respondents rate their chance of dozing in a given situation, from 0 (would never) to 3 (high chance). Total scores range from 0 to 24, with higher scores representing greater sleepiness. Scores over 10 suggest significant daytime sleepiness, and scores over 15 suggest pathological sleepiness associated with conditions like sleep-related breathing disorders or narcolepsy. In the current study, the ESS yielded a Cronbach α of 0.68.

MORNINGNESS-EVENINGNESS QUESTIONNAIRE (MEQ): A shortened version of the MEQ, described elsewhere,[31] was adapted from the original 19-item MEQ[32] and was used in order to reduce participant burden. The MEQ is a self-report measure designed to assess circadian preference, with scores ranging from 6 to 32. Higher scores indicate higher levels of morningness, and lower scores indicate higher levels of eveningness. In the current study, the MEQ yielded a Cronbach α of 0.68.

QUICK INVENTORY OF DEPRESSIVE SYMPTOMS (QIDS): The QIDS[33] is a self-report measure designed to assess depressive symptoms in the past week. The QIDS is a brief version of the 30-item Inventory of Depressive Symptomatology.[34] The measure consists of 16 items on a 4-point Likert scale. Total scores range from 0 to 48, with higher scores indicating greater endorsement of depressive symptoms. A score > 11 indicates clinically significant depression.[33] In the current study, the QIDS yielded a Cronbach α of 0.71.

STATE-TRAIT ANXIETY INVENTORY (STAI): The STAI[35] is a self-report measure designed to assess general anxiety symptoms. The measure consists of 20 statements presented on a 4-point scale. Total scores range from 20 to 80, with higher scores indicating greater endorsement of anxiety symptoms. A score of > 59 indicates clinically significant anxiety.[35] In the current study, the STAI yielded a Cronbach α of 0.92.

PERCEIVED STRESS SCALE (PSS): The PSS[36] is a self-report measure that assesses several domains of stress including unpredictability, lack of control, burden overload, and stressful life circumstances in the past month. The measure consists of 14 items on a 5-point Likert scale ranging from 0 (never) to 4 (very often). Total scores range from 0 to 56, with higher scores representing greater perceived stress. In the current study, the PSS yielded a Cronbach α 0.88.

ALCOHOL USE DISORDERS IDENTIFICATION TEST (AUDIT): The AUDIT[37] is a self-report screening instrument used to detect alcohol consumption that has become harmful to health. The measure consists of 10 items (yes/no and multiple choice) on a 5-point Likert scale ranging from 0 to 4. Total scores range from 0 to 40, with higher scores representing greater harmful alcohol use. A score of ≥ 8 indicates problematic alcohol use. In the current study, the STAI yielded a Cronbach α of 0.83.

MARIJUANA PROBLEMS SCALE (MPS): The MPS[38] is a self-report measure that assesses negative consequences of excessive marijuana use across several domains over the past 90 days. The measure consists of 20 items on a 3-point scale ranging from 0 (no problem) to 2 (serious problem). Total scores range from 0 to 40, with higher scores representing greater marijuana-related consequences. In the current study, the MPS yielded a Cronbach α of 0.90.

Procedure and Analyses

Participants completed the above measures as part of a larger epidemiological study of health in college students. Following informed consent, participants were asked to complete a sleep diary for 7 days. At the end of the week, they returned the sleep diary to study personnel and then completed all other study questionnaires, to ensure the questionnaires were assessing a similar period to the sleep diaries.

Convergent validity (the degree of concordance between the PSQI and another measure of sleep) was assessed by calculating correlation coefficients between the PSQI and weekly averages from the sleep diary for SL, SE, WASO, and TST. Convergent validity was also assessed by calculating correlation coefficients between the PSQI and the ISI and MFI.

Table 2—Correlations between PSQI components and convergent validity measures.

PSQI Component	Mean (SD)	Sleep Diary				ISI^b	MFI^b
		SOL^a	WASO^a	TST^b	SE^b		
1. Sleep Quality	1.06 (0.61)	0.24**	0.29**	−0.17**	−0.34**	0.56**	0.36**
2. Sleep Latency	1.05 (0.93)	0.70**	0.30**	−0.07*	−0.53**	0.39**	0.18**
3. Sleep Duration	0.47 (0.76)	0.02	0.08*	0.51**	−0.22**	0.24**	0.17**
4. Sleep Efficiency	0.52 (0.80)	0.23**	0.21**	−0.05	−0.32**	0.25**	0.18**
5. Sleep Disturbances	1.16 (0.49)	0.19**	0.37**	−0.02	−0.21**	0.33**	0.22**
6. Sleep Medication	0.24 (0.71)	0.13**	0.15**	0.00	−0.15**	0.22**	0.15**
7. Daytime Dysfunction	1.15 (0.76)	0.08*	0.15**	−0.14**	−0.17**	0.46**	0.44**
Global Score	5.64 (2.79)	0.44**	0.37**	−0.26**	−0.52**	0.63**	0.44**

^{a}Indicates correlation coefficient was calculated using Spearman ρ. ^{b}Indicates correlation coefficient was calculated using Pearson *r*. *p < 0.05, **p < 0.01. PSQI, Pittsburgh Sleep Quality Index; SOL, sleep onset latency; WASO, wake time after sleep onset; TST, total sleep time (i.e., duration); SE, sleep efficiency; ISI, Insomnia Severity Index; MFI, Multidimensional Fatigue Inventory-General Fatigue.

Table 3—Correlations between PSQI components and divergent validity measures.

PSQI Component	Self-Report Psychosocial and Sleep Questionnaires						
	ESS^b	MEQ^b	QIDS^b	STAI^b	PSS^b	AUDIT^a	MPS^a
1. Sleep Quality	0.13**	−0.08*	0.37**	0.32**	0.28**	0.05	0.04
2. Sleep Latency	0.00	−0.10**	0.26**	0.19**	0.12**	0.00	0.03
3. Sleep Duration	0.12**	−0.03	0.13**	0.11**	0.11**	−0.00	−0.02
4. Sleep Efficiency	0.02	−0.05	0.18**	0.15**	0.09*	0.04	0.08
5. Sleep Disturbances	0.10**	0.07	0.36**	0.27**	0.26**	0.00	0.06
6. Sleep Medication	0.04	0.04	0.18**	0.16**	0.11**	0.11**	0.06
7. Daytime Dysfunction	0.32**	−0.13**	0.43**	0.37**	0.35**	0.10**	0.01
Global Score	0.18**	−0.08*	0.48**	0.40**	0.33**	0.08*	0.05

^{a}Indicates correlation coefficient was calculated using Spearman ρ. ^{b}Indicates correlation coefficient was calculated using Pearson *r*. *p < 0.05, **p < 0.01. PSQI, Pittsburgh Sleep Quality Index; QIDS, Quick Inventory of Depressive Symptoms; STAI, State-Trait Anxiety Inventory; PSS, Perceived Stress Scale; AUDIT, Alcohol Use Disorders Identification Test; MPS, Marijuana Problems Scale; ESS, Epworth Sleepiness Scale; MEQ, Morningness-Eveningness Questionnaire.

Divergent validity (the degree of concordance between the PSQI and non-sleep measures) was assessed by calculating correlation coefficients between the PSQI and the ESS, MEQ, QIDS, STAI, PSS, AUDIT, and MPS.

We conducted a one-way ANOVA with post hoc (Tukey HSD) to identify differences between extreme morning (n = 7), middle (n = 389), and extreme evening (n = 88) chronotypes on the PSQI.[31]

Results

As shown in **Table 2**, the PSQI global score mean was 5.64 (SD 2.79). There were no significant gender differences in PSQI global score in this sample (mean difference = 0.31, $t_{837} = 1.40$, $d = 0.10$). Significant correlations (ps < 0.05) were demonstrated between the PSQI global score and sleep diary parameters/convergent validity sleep questionnaires. The PSQI global score demonstrated low to moderate correlations with all sleep diary parameters (*r*s = 0.26–0.52), the ISI (*r* = 0.63), and the MFI (*r* = 0.44). The PSQI components demonstrated low to moderate correlations with the analogous sleep diary parameters, ranging from *r* = 0.32 (sleep efficiency) to *r* = 0.70 (sleep latency). The PSQI components demonstrated mostly low correlations with the ISI (*r*s = 0.22–0.56) and the MFI (*r*s = 0.15–0.36).

See **Table 3** for divergent validity calculations. The PSQI global score correlated most with the QIDS (*r* = 0.48) and least with the MPS (*r* = 0.05). The PSQI components demonstrated consistently low correlations with the ESS (*r*s = 0.00–0.32), MEQ (*r*s = 0.03–0.13), QIDS (*r*s = 0.13–0.43), STAI (*r*s = 0.11–0.37), and PSS (*r*s = 0.11–0.35), suggesting some overlap between the PSQI components and non-analogous sleep measures as well as the psychosocial factors of depression, anxiety, and perceived stress. The AUDIT (*r*s = 0.00–0.10) and the MPS (*r*s = 0.01–0.08) demonstrated little to no correlation with the PSQI components or global score, suggesting minimal overlap between the PSQI and substance abuse and good divergent validity with these constructs.

The omnibus ANOVA comparing chronotypes on the PSQI global score was significant, $F_{2,\,481} = 3.34$, p = 0.036. Post hoc tests revealed extreme evening chronotypes (mean = 6.27) had significantly higher scores on the PSQI than middle types (mean = 5.44), p = 0.034, Cohen's *d* = 0.30, but morning types (mean = 6.29) did not differ significantly from middle (*d* = 0.27) or evening types (*d* < 0.01; ps > 0.05). Notably, effect sizes for both extreme types compared to middle types were similar and it is likely that a statistically significant difference was not found between morning and middle types due to small cell sizes.

Table 4—Calculation of true and false negatives and positives.

		DSM-5 Insomnia	
		No	**Yes**
PSQI	**No**	True Negatives (TN)	False Negatives (FN)
	Yes	False Positives (FP)	True Positives (TP)

STUDY II

The aim of Study II was to extend the evaluation of the psychometric properties of the PSQI to a clinical sample of undergraduates diagnosed with insomnia and undergraduate normal sleepers in order to assess the diagnostic accuracy of the PSQI as an insomnia screening tool.

Methods

Sample

A total of 153 healthy college students (71 with insomnia, 82 normal sleepers) volunteered to participate in the parent study in exchange for $105. Six participants were excluded from the current analyses for missing data on the PSQI. The final sample size for the current analyses was n = 147 (87 female) with an average age of 20.2 years (SD = 2.6). Sample characteristics are presented in **Table 1**. The racial and ethnic diversity was as follows: 65.3% Non-Hispanic White, 11.6% African American/Black, 3.4% Asian/Pacific Islander, 19.8% Biracial or Other, and 29.3% Hispanic. This study was approved by the university institutional review board, and informed consent was obtained for all participants.

Measurement

PSQI: The PSQI (described above) used in Study II was a 19-item self-rated retrospective questionnaire of the past month. Cronbach α in Study II was 0.87.

STRUCTURED CLINICAL INTERVIEW SCHEDULE FOR DSM-5 SLEEP DISORDERS: Participants were interviewed by trained master's level psychology student interviewers using a Structured Clinical Interview Schedule to assess for current sleep disorders as defined by the Diagnostic and Statistical Manual, fifth edition (DSM-5).[39] Interviewers were supervised by a licensed clinical psychologist board certified in both Sleep Medicine and Behavioral Sleep Medicine (DJT). For the purposes of this study, only participants who were classified as either meeting criteria for Insomnia Disorder or not meeting criteria for any sleep disorder ("normal sleepers") were included for analyses.

Procedure and Analyses

All participants completed a variety of questionnaires including the PSQI and then underwent structured clinical interviews for sleep disorders to ensure insomnia diagnoses, and psychiatric interviews, physical examinations, and drug testing to ensure comorbid health problems were not potential confounds.

Table 5—Sensitivity, specificity, and Youden's index of PSQI compared to DSM-5 insomnia disorder diagnosis.

PSQI Cutoff	Sensitivity	Specificity	Youden's Index (*J*)
≥ 4	1.00	0.73	0.73
≥ 5*	1.00	0.92	0.92
≥ 6	0.99	1.00	0.99
≥ 7	0.94	1.00	0.94

*Original validation study cutoff value.[1] PSQI, Pittsburgh Sleep Quality Index.

The ability of the PSQI to differentiate between individuals with and without insomnia was calculated using analyses of sensitivity, specificity, the receiver operating characteristic (ROC) curve, and Youden Index. These analyses were performed using DSM-5 insomnia diagnosis as the "gold standard" against which the PSQI was compared.[39] DSM-5 diagnostic criteria for Insomnia Disorder includes ≥ 3 nights per week of difficulties with sleep onset latency, wake after sleep onset, or terminal wakefulness, a daytime complaint (e.g., fatigue, difficulty concentrating), significant distress, and a duration ≥ three months.[39]

See **Table 4** for explanation of true negatives (TN), false negatives (FN), false positives (FP), and true positives (TP). Sensitivity, the probability the PSQI can correctly detect individuals with insomnia, was calculated by TP / (TP + FN). Specificity, the probability the PSQI can correctly identify individuals without insomnia, was calculated by TN / (TN + FP). The ROC curve plots sensitivity (y axis) and 1 – specificity (x axis) and produces the area under the curve (AUC) metric to help assess a test's diagnostic accuracy.[40] The Youden index is a summary statistic of the ROC curve used to determine the optimal cutoff value by optimizing the test's differentiating ability when sensitivity and specificity are valued equally.[41] Analyses were performed using SPSS 20 (IBM).

Results

Table 5 presents the sensitivity and specificity of the PSQI and Youden Index scores at various cutoff points compared to the "gold standard" of DSM-5 Insomnia Disorder diagnosis. The AUC was 0.999, suggesting the PSQI has extremely high sensitivity and specificity in its detection of insomnia diagnosis. The Youden Index score was maximized (0.99) at a cutoff score of 6, which is slightly higher than the original validation study cutoff of 5.[1]

DISCUSSION

Examination of the relationship between the PSQI and sleep diaries/other self-report questionnaires suggests the PSQI is only moderately related to most sleep diary components and measures of insomnia and fatigue. Examination of the relationship between the PSQI and other sleep and psychosocial factors (i.e., sleepiness, circadian preference, depression, anxiety, perceived stress, alcohol abuse, marijuana abuse) suggests although the PSQI has good divergent validity with the Morningness-Eveningness Scale, Alcohol Use Disorders Identification Test and the Marijuana Problems Scale, it substantially overlaps in measurement with the Epworth Sleepiness Scale, Quick Inventory of Depressive Symptoms, the State-Trait Anxiety Inventory, and the Perceived Stress Scale. The results of this study indicate PSQI component scores may provide useful information about sleep in college students, although caution should be used when interpreting the PSQI global score as a measure of sleep disturbance because of the broad, multifaceted nature of the components that make up this score. Although the sleep diary and the PSQI purport to measure similar dimensions of sleep, correlations between analogous metrics on these measures indicate only low to moderate concordance. Results examining sensitivity, specificity, and diagnostic accuracy of Insomnia Disorder by the PSQI suggest this measure performs well as an insomnia screening tool in this population, but users are cautioned that this measure should not be used as a diagnostic tool for insomnia due to the substantial overlap with other constructs.

The average PSQI global score for the sample in Study I was mean = 5.64 (SD = 2.79). Compared to the original validation study, this was higher than controls (mean = 2.67, SD = 1.70), slightly lower than patients with disorders of excessive sleepiness (mean = 6.53, SD = 2.98) and significantly lower than patients with disorders of initiating and maintaining sleep (mean = 10.38, SD = 4.57).[1] The average PSQI global score mean in Study I was likely higher than controls and lower than individuals reporting sleep complaints because this was a sample that contained individuals with insomnia and other sleep disorders as well as individuals reporting no sleep symptoms. The average PSQI global score in Study II for individuals with insomnia (mean = 10.65, SD = 2.79) and for controls (mean = 2.63, SD = 1.29) were similar to the original validation study (see above). This suggests the PSQI performed similarly in this population compared to the middle-aged adults examined in the original validation study.[1]

Although the correlations of the PSQI with analogous sleep diary variables were significant, the magnitudes were at best moderate.[42] They were, however, somewhat higher than those found by another study who reported similar values in young adults (e.g., sleep disturbances $r = 0.07$, sleep efficiency $r = 0.01$).[12] This lack of strong correlation at first glance appears concerning, considering both measures purport to assess aspects of the same construct (e.g., sleep quality). However, the PSQI is a broad measure that assesses many dimensions of sleep (e.g., "sleep medication"; "daytime dysfunction") not captured by the sleep diary. Furthermore, the PSQI uses single-time-point retrospective estimate, which has known issues of bias such as systematic overestimation.[43] Finally, previous studies have found younger individuals' PSQI global scores do not correlate as highly with analogous sleep diary metrics compared to older individuals.[12]

The PSQI was compared to validated measures of other thematically related sleep constructs in order to determine if it demonstrated convergent validity. The ISI demonstrated moderate convergent validity with the PSQI global score ($r = 0.63$) and Subjective Sleep Quality component ($r = 0.56$), and low convergent validity (rs = 0.22–0.46) with the other PSQI components. Therefore, if one was interested in specifically assessing insomnia, it would be wiser to choose the smaller 7-item ISI over the 19-item PSQI. Similarly, the MFI demonstrated low convergent validity with the PSQI global score ($r = 0.44$), Sleep Quality ($r = 0.36$), and Daytime Dysfunction ($r = 0.44$) components. This would suggest the PSQI was not very sensitive to complaints of fatigue. Thus, one would be better served using fatigue specific instruments to assess this domain.

The ESS demonstrated low correlations with the PSQI components, the highest being the Daytime Dysfunction component ($r = 0.32$). It appears the PSQI is slightly less sensitive to conditions that produce excess daytime sleepiness, such as narcolepsy and sleep-related breathing disorders, compared to insomnia or fatigue. This is unsurprising, given that sleepiness and fatigue are distinct constructs and a large degree of overlap would not be expected.[30] For instance, previous studies have shown that insomnia, which is arguably primarily what the PSQI is measuring, is highly correlated with daytime fatigue but not with excessive daytime sleepiness.[44]

Unsurprisingly, the global PSQI had very low correlations with the MEQ, suggesting the PSQI does not pick up on non-pathological circadian rhythm differences. Contrasting extreme chronotypes with middle chronotypes revealed moderate differences in subjective sleep quality. Specifically, this relationship appears to be U-shaped, with middle chronotypes reporting the best subjective sleep quality and extreme types reporting worse approximately equivalent sleep quality. However, the morning-type results should be interpreted cautiously as, consistent with previous literature, the prevalence of extreme morning types was extremely small in this population.

The PSQI had significant overlap with the QIDS, STAI, and PSS, suggesting that it is sensitive to depression, anxiety, and perceived stress. This is unsurprising given previous findings of high correlations between the PSQI and psychological symptoms.[21] The QIDS has several questions that assess sleep, but even after these items were removed, correlations remained moderate. It is likely that an underlying factor (e.g., hyperarousal) produces endorsement of both sleep and psychosocial complaints, as these factors are known to frequently present in comorbid fashion and interact with one another.[45] Use of the PSQI as a screening tool should come with the awareness that this relationship exists, and that high endorsement on the PSQI may reflect high psychological distress. Conversely, the PSQI demonstrated low correlations with the AUDIT and the MPS, suggesting substance abuse is not highly related to sleep complaints as measured by the PSQI in this population. This may be due in part to the relative ubiquitous use of alcohol and marijuana in this population as compared to other groups.[46]

The results of Study II demonstrated the PSQI very accurately identified individuals with and without insomnia. The PSQI performed substantially better at identifying DSM-5 diagnosed insomnia in this U.S. college population than it did in a previous study attempting to identify DSM-IV diagnosed insomnia in a sample of Nigerian college students.[13] This may be because the current study performed extensive assessment procedures (e.g., self-report measures, clinical interviews, history and physicals, lab assessments) to select otherwise healthy people (e.g., no other underlying sleep, psychiatric, or medical disorders) with and without insomnia, whereas the other study did not rule out comorbidities. If Study II had not used such rigorous assessments and had not excluded people with symptoms of underlying sleep disorders (e.g., apnea, circadian rhythm sleep disorders, sleep deprivation), sleep medication usage, or subclinical insomnia, the PSQI likely would not have performed as well.

The results of Study II indicate the optimal sensitivity and specificity for insomnia detection was at a cutoff score of 6. This cutoff could be adjusted depending on the intended use of the PSQI, as sensitivity and specificity are still very high at surrounding cutoff scores. For example, those seeking to screen for any individuals with potential insomnia complaints may choose the original validation study cutoff of 5 to maximize sensitivity and sacrifice a small amount of specificity.

The PSQI has both strengths and limitations in college students. The PSQI appears to be best able to detect insomnia symptoms and fatigue complaints. The PSQI serves as an accurate diagnostic screening tool for insomnia in this population (healthy college students with no comorbidities). However, the PSQI should not be used alone to diagnose insomnia—a comprehensive sleep evaluation including clinical interview is optimal to successfully assign sleep disorder diagnoses. Previous studies[47,48] have demonstrated PSQI scores are more biased in psychiatric populations compared to non-psychiatric populations, and urge cautious interpretation when using the PSQI in individuals with a history of mental illness. Further, the PSQI is less sensitive to daytime sleepiness complaints and circadian preference, which can significantly impact assessment and treatment of insomnia and other potentially comorbid sleep disorders. Finally, the PSQI detects depression, anxiety, and perceived stress, and should therefore be used in conjunction with measures of these constructs as part of a holistic assessment.

The results of this study are limited by several factors. The version of the PSQI used in this study asked about the previous 7 days, whereas the originally published PSQI asked about the previous month.[1] This time period was chosen to mirror the time period for sleep diary data collection and to mimic the short assessment periods commonly found in physicians' and other general practice. Although this limitation does impact the interpretation of results because longer periods of assessment have been demonstrated to provide greater accuracy with respect to highly variable factors like sleep, it appears that a week duration may be sufficient to capture this variability.[49,50] Additionally, with regard to time frame, the various measures used in this study did not all assess similar time periods. This affects the comparability of these measures and may understate the relationships that exist between them. Second, the samples included for these studies were from a single institution, and may not be generalizable to students at institutions in other regions or with other racial/ethnic and/or socioeconomic make-ups. Third, the sample used in Study II was healthy and did not have other diagnoses, which limits generalizability to individuals with comorbid sleep, psychological, and medical conditions. It is likely that the diagnostic accuracy of the PSQI would be substantially reduced among individuals with these comorbidities.

The heterogeneous nature of sleep complaints in this population indicates a need for precise and thorough measurement tools. Although college students may report sleep habits that appear disordered compared to middle-aged adults, these habits may be more normative in this population. The PSQI may be a good screening tool, but is not a substitute for sleep disorder diagnoses established through a clinical interview performed by an experienced sleep clinician. The reasons for this are multi-fold, but are primarily because (a) the PSQI assesses many different sleep disorders domains (e.g., sleep duration, sleep disturbances, use of sleep medications, daytime dysfunction), (b) the PSQI is moderately sensitive to complaints of insomnia, fatigue, and subjective distress (i.e., depression, anxiety, and perceived stress), and (c) an interview can disentangle the possible effects of negative affect bias on subjective sleep distress reporting.

If the PSQI is to be used in a college population as an insomnia screening tool, a cutoff score ≥ 6 is recommended in order to maintain high sensitivity of the test and maximize the diagnostic accuracy. The PSQI may not accurately reflect daytime sleepiness, circadian preference, or alcohol or marijuana substance abuse problems and it is recommended that other validated tools should be used to assess the latter constructs. Users of the PSQI in populations with a history of psychiatric illness should exercise caution in interpretation. Finally, future research should examine or develop measurement tools that account for college student's unique sleep habits and should examine the psychometrics of the PSQI in populations with comorbid sleep and psychological conditions.

ABBREVIATIONS

AUC, area under the curve
AUDIT, Alcohol Use Disorders Identification Test
DSM, Diagnostic and Statistical Manual
ESS, Epworth Sleepiness Scale
FN, false negatives
FP, false positives
ISI, Insomnia Severity Index
MEQ, Morningness-Eveningness Questionnaire
MFI, Multidimensional Fatigue Inventory
MPS, Marijuana Problems Scale
PSQI, Pittsburgh Sleep Quality Index
PSS, Perceived Stress Scale
QIDS, Quick Inventory of Depressive Symptoms
ROC, receiver operating characteristic
SD, standard deviation
SE, sleep efficiency
SL, sleep latency
STAI, State-Trait Anxiety Inventory
TN, true negatives
TP, true positives
TST, total sleep time
WASO, wake after sleep onset

REFERENCES

1. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiat Res 1989;28:193–213.

2. Pallos H, Yamada N, Doi Y, Okawa M. Sleep habits, prevalence and burden of sleep disturbances among Japanese graduate students. Sleep Biol Rhythms 2004;2:37–42.
3. Tzeng JI, Fu Y, Lin C. Validity and reliability of the Taiwanese version of the Pittsburgh Sleep Quality Index in cancer patients. Int J Nurs Stud 2012;49:102–8.
4. Bertolazi AN, Fagondes SC, Hoff LS, et al. Validation of the Brazilian Portuguese version of the Pittsburgh Sleep Quality Index. Sleep Med 2011;12:70–5.
5. Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. J Psychosom Res 2002;53:737–40.
6. Fichtenberg NL, Putnam SH, Mann NR, Zafonte RD, Millard AE. Insomnia screening in postacute traumatic brain injury: utility and validity of the Pittsburgh Sleep Quality Index. Am J Phys Med Rehab 2001;80:339–45.
7. Beck SL, Schwartz AL, Towsley G, Dudley W, Barsevick A. Psychometric evaluation of the Pittsburgh sleep quality index in cancer patients. J Pain Symptom Manag 2004;27:140–8.
8. Carpenter JS, Andrykowski MA. Psychometric evaluation of the Pittsburgh Sleep Quality Index. J Psychosom Res 1998;45:5–13.
9. Burkhalter H, Sereika SM, Engberg S, Wirz-Justice A, Steiger J, De Geest S. Structure validity of the Pittsburgh Sleep Quality Index in renal transplant recipients: a confirmatory factor analysis. Sleep Biol Rhythms 2010;8:274–81.
10. Gentili A, Werner DK, Kuchibhatla M, Edinger JD. Test-retest reliability of the Pittsburgh sleep quality index in nursing home residents. J Am Geriatr Soc 1995;43:1317–8.
11. Skouteris H, Germano C, Wertheim EH, Paxton SJ, Milgrom J. Sleep quality and depression during pregnancy: a prospective study. J Sleep Res 2008;17:217–20.
12. Grandner MA, Kripke DF, Yoon I, Youngstedt SD. Criterion validity of the Pittsburgh Sleep Quality Index: investigation in a non-clinical sample. Sleep Biol Rhythms 2006;4:129–36.
13. Aloba OO, Adewuya AO, Ola BA, Mapayi BM. Validity of the Pittsburgh Sleep Quality Index (PSQI) among Nigerian university students. Sleep Med 2007;8:266–70.
14. Yang CM, Wu CH, Hsieh MH, Liu MH, Lu FH. Coping with sleep disturbances among young adults: a survey of first-year college students in Taiwan. Behav Med 2003;29:133–8.
15. Taylor DJ, Bramoweth AD, Grieser EA, Tatum JI, Roane BM. Epidemiology of insomnia in college students: relationship with mental health, quality of life, and substance use difficulties. Behav Ther 2013;44:339–48.
16. Lund HG, Reider BD, Whiting AB, Prichard JR. Sleep patterns and predictors of disturbed sleep in a large population of college students. J Adolescent Health 2010;46:124–32.
17. Taylor DJ, Gardner CE, Bramoweth AD, et al. Insomnia and mental health in college students. Behav Sleep Med 2011;9:107–16.
18. Lee S, Wuertz C, Rogers RL, Chen Y. Stress and sleep disturbances in female college students. Am J Health Behav 2013;37:851–8.
19. Pilcher JJ, Huffcutt AJ. Effects of sleep deprivation on performance: a meta-analysis. Sleep 1996;19:318–26.
20. Kelly WE, Kelly KE, Clanton RC. The relationship between sleep length and grade-point average among college students. Coll Stud J 2001;35:84–6.
21. Buysse DJ, Hall ML, Strollo PJ, et al. Relationships between the Pittsburgh Sleep Quality Index (PSQI), Epworth Sleepiness Scale (ESS), and clinical/polysomnographic measures in a community sample. J Clin Sleep Med 2008;4:563–71.
22. Taylor DJ, Bramoweth AD, Grieser EA, Tatum JI, Roane BM. Epidemiology of insomnia in college students: relationship with mental health, quality of life, and substance use difficulties. Behav Ther 2013;44:339–48.
23. Lichstein KL, Stone KC, Donaldson J, et al. Actigraphy validation with insomnia. Sleep 2006;29:232–9.
24. Vallières A, Morin CM. Actigraphy in the assessment of insomnia. Sleep 2003;26:902–6.
25. Carney CE, Buysse DJ, Ancoli-Israel S, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 2012;35:287–302.
26. Taylor DJ, Vatthauer KE, Bramoweth AD, Ruggero C, Roane B. The role of sleep in predicting college academic performance: is it a unique predictor? Behav Sleep Med 2013;11:159–72.
27. Taylor DJ, Zimmerman MR, Gardner CE, et al. A pilot randomized controlled trial of the effects of cognitive-behavioral therapy for insomnia on sleep and daytime functioning in college students. Behav Ther 2014;45:376–89.
28. Bastien CH, Vallieres A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001;2:297–307.
29. Smets EM, Garssen B, Bonke B, De Haes JC. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res 1995;39:315–25.
30. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991;14:540–5.
31. Taylor DJ, Clay KC, Bramoweth AD, Sethi K, Roane BM. Circadian phase preference in college students: relationships with psychological functioning and academics. Chronobiol Int 2011;28:541–7.
32. Horne JA, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiology 1976;4:97–110.
33. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiat 2003;54:573–83.
34. Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The inventory of depressive symptomatology (IDS): psychometric properties. Psychol Med 1996;26:477–86.
35. Spielberger CD, Gorsuch RL, Lushene R. The state-trait anxiety inventory (test manual). Palo Alto, CA: Consulting Psychologists, 1970.
36. Cohen S. Perceived stress in a probability sample of the United States In: Spacapan S, Oskamp S, eds. Perceived stress in a probability sample of the United States. City: Sage, 1988:31–67.
37. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993;88:791–804.
38. Stephens RS, Wertz JS, Roffman RA. Predictors of marijuana treatment outcomes: the role of self-efficacy. J Subst Abuse 1993;5:341–54.
39. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Arlington, VA: American Psychiatric Publishing, 2013.
40. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993;39:561–77.
41. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32–5.
42. Cohen J. Statistical power analysis for the behavioral sciences. Psychology Press, 1988.
43. Johnson DC, Polusny MA, Erbes CR, et al. Development and initial validation of the Response to Stressful Experiences Scale. Mil Med 2011;176:161–9.
44. Riedel BW, Lichstein KL. Insomnia and daytime functioning. Sleep Med Rev 2000;3:277–98.
45. Taylor DJ, Lichstein KL, Durrence HH, Riedel BW, Bush AJ. Epidemiology of insomnia, depression, and anxiety. Sleep 2005;28:1457–64.
46. O'Malley PM, Johnston LD. Epidemiology of alcohol and other drug use among American college students. J Stud Alcohol Supp 2002:23–39.
47. Hartmann JA, Carney CE, Lachowski A, Edinger JD. Exploring the construct of subjective sleep quality in patients with insomnia. J Clin Psychiatry 2015;76:e768–73.
48. Calhoun PS, Wiley M, Dennis MF, Means MK, Edinger JD, Beckham JC. Objective evidence of sleep disturbance in women with posttraumatic stress disorder. J Trauma Stress 2007;20:1009–18.
49. Knutson KL, Rathouz PJ, Yan LL, Liu K, Lauderdale DS. Intra-individual daily and yearly variability in actigraphically recorded sleep measures: the CARDIA study. Sleep 2007;30:793–6.
50. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. Sleep 2006;29:1155–73.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication March, 2015
Submitted in final revised form March, 2016
Accepted for publication April, 2016

Address correspondence to: Daniel J. Taylor, PhD, ABSM, CBSM, Professor, University of North Texas, Department of Psychology, 1155 Union Circle #311280, Denton, Texas 76203-5017; Tel: (940) 565-2655; Fax: (940) 565-4682; Email: Daniel.Taylor@unt.edu

DISCLOSURE STATEMENT

This was not an industry supported study. This study was supported by a grant from the University of North Texas (G69250) and from the National Institute of Allergy and Infectious Diseases (1R15AI085558-01A2). Dr. Roane has consulted for Johnson & Johnson. The other authors have indicated no financial conflicts of interest. The study was performed at the University of North Texas, Denton, TX.