https://doi.org/10.5664/jcsm.9426

JCSM | Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

The Hypersomnia Severity Index: reliability, construct, and criterion validity in a clinical sample of patients with sleep disorders

Julio Fernandez-Mendoza, PhD, DBSM[1]; Kristina Puzino, PhD[1]; Gregory Amatrudo, PsyD[1]; Elizaveta Bourchtein, PhD[1]; Susan L. Calhoun, PhD, ABPP[1]; David T. Plante, MD, PhD[2]; Kate Kaplan, PhD, DBSM[3]

[1]Sleep Research and Treatment Center, Department of Psychiatry and Behavioral Health, Penn State Health Milton S. Hershey Medical Center, Penn State College of Medicine, Hershey, Pennsylvania; [2]Department of Psychiatry, University of Wisconsin-Madison, Madison, Wisconsin; [3]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Study Objectives: The Hypersomnia Severity Index (HSI) was designed to assess the severity and impairment of hypersomnolence and has been validated in persons with psychiatric disorders. Little is known about its psychometric properties in clinical samples of patients with sleep disorders.

Methods: One hundred fifty-eight patients (aged 44.1 ± 16.4 years, 29.1% male, 19.6% racial/ethnic minority) evaluated at the Behavioral Sleep Medicine program of the Penn State Health Sleep Research and Treatment Center completed the HSI and other patient-reported outcomes. We examined the HSI's reliability and factorial, construct, and criterion validity.

Results: The HSI showed satisfactory internal consistency (α = 0.79). A 2-factor structure, reflecting symptoms (HSI-S) and impairment, explained 56.2% of the variance. Convergent validity with the Epworth Sleepiness Scale was optimal (r = .65) but greater for HSI-S (r = .69) than for impairment (r = .39). Divergent validity was optimal for HSI-S against unrelated measures of sleep effort, reactivity, and incompatible behaviors ($r \leq .02$). Construct validity showed higher scores in patients with central disorders of hypersomnolence and lower scores in patients with chronic insomnia disorder compared to those with other sleep disorders; however, these divergent scores were primarily driven by HSI-S rather than impairment. Criterion validity showed that an HSI-S cutoff score ≥ 8 provided the best balance in sensitivity/specificity (0.82/0.78) to identify central disorders of hypersomnolence (area under the curve, 0.85).

Conclusions: The HSI shows satisfactory indices of reliability and validity in a clinical patient sample. Its construct and criterion validity are supported by its divergent association with other patient-reported outcomes and central disorders of hypersomnolence vs chronic insomnia disorder diagnoses and the adequate sensitivity/specificity of its HSI-S cutoff score to reliably identify central disorders of hypersomnolence.

Keywords: hypersomnia, patient-reported outcomes, severity, screening

Citation: Fernandez-Mendoza J, Puzino K, Amatrudo G, et al. The Hypersomnia Severity Index: reliability, construct, and criterion validity in a clinical sample with sleep disorders. *J Clin Sleep Med.* 2021;17(11):2249–2256.

BRIEF SUMMARY

Current Knowledge/Study Rationale: Developing patient-reported outcome measures for central disorders of hypersomnolence that adequately assess essential symptoms and associated daytime impact is critical. The Hypersomnia Severity Index was designed to assess the severity, distress, and impairment of hypersomnolence and has been validated in individuals with mood disorders; however, its psychometric properties in clinical samples of patients with sleep disorders have remained unexplored.

Study Impact: The present study supports that the Hypersomnia Severity Index assesses 2 factors, hypersomnia symptoms and associated distress and impairment, with optimal reliability. Its construct and criterion validity are also supported by its convergent and divergent association with other patient-reported outcomes and subtypes of central disorders of hypersomnolence and its adequate sensitivity/specificity to identify central disorders of hypersomnolence.

INTRODUCTION

Excessive daytime sleepiness and excessive need to sleep (ie, hypersomnolence) are among the most prevalent sleep-related concerns in the general population[1] and the chief concerns in clinical samples of patients presenting to sleep centers.[2] Cross-sectional and longitudinal studies have shown that a myriad of psychiatric, cardiometabolic, and sleep disorders, among others, are risk factors for hypersomnolence,[3–5] which is in itself associated with significant personal and occupational sequelae and public safety hazards.[6–8]

The central disorders of hypersomnolence (CDH) are a group of disorders manifesting as an irrepressible need to sleep, daytime sleep episodes, and daytime impairment that cannot be attributed solely to disrupted nocturnal sleep or a circadian misalignment. These disorders include narcolepsy, idiopathic hypersomnia, hypersomnia because of a medical or neurologic disorder, or hypersomnia associated with psychiatric disorders. The current standard assessment for patients with suspected CDH is an in-laboratory, nocturnal polysomnography followed by a daytime Multiple Sleep Latency Test, with all but a few diagnoses of CDH requiring such testing. This in-laboratory evaluation assists in identifying patients with increased physiological sleep propensity and in the differential diagnosis of narcolepsy and other forms of CDH, including idiopathic hypersomnia. Despite their valuable diagnostic purposes, neither polysomnography nor

the Multiple Sleep Latency Test provides a measure of the severity or impairment associated with hypersomnolence or CDH from the patient's perspective. It is likely that a combination of such objective sleep measures and patient-centered outcomes may provide a better screening, diagnosis, and severity assessment of hypersomnolence.

Several self-reported measures have been developed to capture individual features of hypersomnolence. For instance, the Epworth Sleepiness Scale (ESS)[9] is a global measure of self-reported sleep propensity across multiple disorders but is not necessarily a measure of the severity and daytime impact of hypersomnolence and/or CDH. Recently, patient-reported outcomes specifically designed to assess the severity of hypersomnolence and CDH have been developed.[10,11] Kaplan and colleagues[10] developed the Hypersomnia Severity Index (HSI) to assess essential features of hypersomnia (excessive sleepiness and need to sleep) together with associated distress and impairment. The initial psychometric properties of the HSI were examined in a sample of 381 undergraduate students (mean age 21 years, 55.0% female, 49.3% White), 89 patients with hypersomnia associated with bipolar disorder (mean age 35 years, 61.8% female, 70.1% White), and 21 patients with hypersomnia-associated major depressive disorder (mean age 28 years, 81.0% female, 90.0% White). The study showed optimal reliability and a 2-factor structure, representing hypersomnia symptoms (HSI-S) and distress/impairment (HSI-I) and adequate convergent validity with the ESS and other sleep-related and mood-related patient-reported outcomes.[10] The HSI is thus a promising measure for assessing the presence and severity of hypersomnolence.

Given the limited data on the HSI in clinical samples of patients with sleep disorders, including CDH, the goal of the present analysis was to expand upon the previous study[10] by examining the reliability, factor structure, and construct and criterion validity of the HSI as administered to a clinical sample of patients with diverse sleep disorders. Specifically, we aimed to replicate the factorial validity of the HSI and its suggested HSI-S and HSI-I factors. We also planned to test whether the scale had robust convergent validity with other measures of theoretically related constructs and discriminant validity with theoretically unrelated constructs. In addition, we aimed to test the construct validity of the HSI and its proposed factors by examining differences between 3 diagnostic groups. We hypothesized that HSI scores would be higher in the CDH group than in a chronic insomnia disorder (CID) or other sleep disorders (OSD) group. On an exploratory basis, we also aimed to evaluate differences between CDH subgroups, namely, those with psychiatric hypersomnia, idiopathic hypersomnia and narcolepsy. Finally, we intended to examine the criterion validity of the HSI in terms of its concurrent validity with the presence of CDH to inform the development of optimal cutoff scores to identify clinically significant hypersomnia among patients with sleep disorders.

METHODS

Patient sample

A retrospective chart review of the electronic medical record was conducted for the period March 1, 2014–February 22, 2019 to identify patients who completed the HSI as part of their evaluation at the behavioral sleep medicine program of the Penn State Sleep Research and Treatment Center (Hershey, PA). A total of 445 patients had undergone a behavioral sleep medicine evaluation during the retrospective 5-year review period. The HSI was first administered to a given patient starting July 1, 2015, but it did not become part of the standard testing packet until April 4, 2017. A total of 158 patients had complete data on the HSI, and the vast majority (79.8%, n = 126) of the evaluations were completed during the April 4, 2017–February 22, 2019 period. Study procedures were approved by the Penn State College of Medicine Institutional Review Board (IRB-00009522).

Clinical data

All 158 patients identified had undergone a thorough clinical history via a semistructured clinical interview and comprehensive testing that included validated patient-reported scales. Participants' demographic information such as age, sex, race, body mass index, and blood pressure at the time of the diagnostic visit were retrieved from the electronic medical record. All sleep disorder diagnoses were assigned as per the *International Classification of Sleep Disorders,* third edition criteria[12] by board-certified sleep physicians or sleep psychologists. In the present study, patients were identified as having a CDH (n = 39) if they had received a diagnosis of type 1 or 2 narcolepsy (n = 10), idiopathic hypersomnia (n = 15), or hypersomnia associated with a psychiatric (ie, mood) disorder (n = 14). For the purpose of this study, patients were identified as having CID (n = 73) if they had received such a diagnosis in the absence of any other comorbid sleep or circadian disorder. Finally, patients were identified as having OSD (n = 46) if they had received a diagnosis of a sleep-related breathing disorder (eg, obstructive sleep apnea), circadian rhythm sleep-wake disorder (eg, delayed sleep-wake phase disorder), parasomnia (ie, nonrapid eye movement sleep– and rapid eye movement sleep–related) or sleep-related movement disorder (eg, restless legs syndrome). Data regarding patient-reported scales were retrieved, and those assessing hypersomnolence, sleepiness, fatigue, depression, sleep reactivity, sleep effort, and sleep-incompatible behaviors were the focus of this psychometric study.

The HSI comprises 9 items, each rated on a Likert scale from (0) not at all–(4) very much, producing a total score ranging from 0–36. It has shown optimal internal consistency (Cronbach's α = 0.84) and convergent validity (eg, r = .44 with the ESS) in samples of undergraduate students and persons with mood disorders.[10] Previous factorial validity has identified HSI items 1a (sleeping too much at night), 1b (having difficulty waking up in the morning or from naps), 1c (sleeping during the day), 1d (feeling sleepy during the daytime), and 6 (having "sleep attacks"/unintended sleep in inappropriate situations) as those constituting the HSI-S factor and items 2 (satisfaction/dissatisfaction with sleep pattern), 3 (interference with daily functioning), 4 (how noticeable the sleep problem is to others), and 5 (worry/distress about sleep problem) as those constituting the HSI-I factor. Higher HSI total scores (HSI-T) indicate greater severity of hypersomnia symptoms and associated distress/impairment. An HSI-T of ≥ 10 has been suggested to discriminate between undergraduate students

and individuals with hypersomnia associated with a mood disorder, with higher scores in the latter group.[10]

Convergent measures

ESS

The ESS is an 8-item self-report measure of excessive daytime sleepiness.[9] Items assess the propensity for falling asleep in common daytime situations, yielding a composite score of sleepiness severity. The 8 items on the ESS are rated on a Likert scale from 0 = no chance to 3 = high chance, producing a total score ranging from 0–24. The ESS has shown good internal consistency and high test-retest reliability.[9] A total score > 10 has been suggested to identify excessive daytime sleepiness. The ESS was used in the current study as a measure of daytime sleepiness that theoretically converges with the HSI.

Flinders Fatigue Scale

The Flinders Fatigue Scale (FFS) is a 7-item scale that asks about the experience of fatigue in the previous 2 weeks.[13] Responses range from 0 = not at all to 4 = extremely. Total scores range from 0–31, with higher scores indicating higher levels of experienced fatigue. The FFS has shown good reliability and validity, with a total score of ≥ 16 identifying moderate to severe daytime fatigue.[13] The FFS was used in the current study as a measure of daytime fatigue that theoretically converges with the HSI.

Depression Anxiety Stress Scale

The Depression Anxiety Stress Scale is a 42-item scale used to assess negative mood states over the past week.[14] Items are rated on a 4-point Likert scale (0 = does not apply to me at all to 3 = applied to me very much, or most of the time), producing total scores for each subscale ranging from 0–42, where higher scores indicate a greater severity of negative symptoms. Scores for the Depression Anxiety Stress Scale-depression (DASS-D) subscale were used in the current study, with a score of ≥ 14 indicating moderate to severe depressive symptoms.[15] The DASS-D was used in the current study as a measure of depression that theoretically converges with the HSI.

Divergent measures

Ford Insomnia Response to Stress Test

The Ford Insomnia Response to Stress Test (FIRST) is a 9-item scale administered as a measure of sleep reactivity.[16] The FIRST requires the individual to rate the likelihood of having sleep disruption in association with specific and common stressful events or periods of stress occurring during the day or evening, on a 4-point scale (1 = not likely, 2 = somewhat likely, 3 = moderately likely, 4 = very likely). The FIRST has shown adequate reliability and validity, including high test-retest reliability (r = .92).[16] Higher scores on the FIRST are indicative of greater trait predisposition to stress-related sleep disturbance, with a score of ≥ 16 indicating high sleep reactivity.[17] The FIRST was used in the current study as a measure of stress-related sleep reactivity that theoretically diverges from the HSI.

Glasgow Sleep Effort Scale

The Glasgow Sleep Effort Scale (GSES) is a 7-item self-report questionnaire reflecting the perceived controllability of sleep.[18] Total scores range from 0–14, with higher scores indicating greater sleep effort over the past week. The GSES has shown moderate to high scale reliability and strong sensitivity and specificity, successfully discriminating good sleepers from individuals with CID.[19] The GSES was used in the current study as a measure of sleep effort that theoretically diverges from the HSI.

Sleep Hygiene Practice Scale

The Sleep Hygiene Practice Scale (SHPS) is a 30-item scale that assesses the practice of daily living activities and sleep habits that may have negative impacts on sleep.[20] The Sleep Hygiene Practice Scale adopts a multifactorial structure to measure sleep hygiene by including environmental factors and behavioral practices that may impact sleep. Individuals rate how frequently they engage in the behavioral practices or how much specific environmental factors impact their sleep based on a 6-point Likert scale. The scores for the items are summed to generate 5 subscale scores representing different dimensions of sleep hygiene and behaviors. The Sleep Hygiene Practice Scale has shown optimal reliability and factor validity and strong sensitivity and specificity to discriminate good sleepers from individuals with CID.[21] The Sleep Hygiene Practice Scale sleep-incompatible behaviors (SHPS-I) subscale was used in the current study as a measure that theoretically diverges from the HSI.

Statistical analysis

All data were analyzed using SPSS, version 25 (IBM Corp., Armonk, N.Y., USA). All data were examined for distribution, kurtosis, and missing values. All variables were normally distributed.

Internal consistency was examined using Cronbach's alpha. Exploratory factor analysis was conducted to examine the factorial validity of the HSI as it pertained to its purported assessment of essential symptoms and daytime impact of hypersomnolence. The number of factors retained in exploratory factor analysis using principal component analyses with Varimax rotation and Kaiser normalization was determined based on eigenvalues in the scree plot. We chose this method to remain commensurate with previous reports on the HSI.[10] A cutoff value of 0.4 was set for determining the saliency of factor loadings.[22]

To examine construct validity, Pearson correlations were conducted to evaluate the convergence between HSI with theoretically related measures of daytime sleepiness (ESS), fatigue (FFS), and depression (DASS-D) and its divergent validity with theoretically unrelated measures of sleep reactivity (FIRST), sleep effort (GSES), and sleep-incompatible behaviors (SHPS-I). Cohen's conventions were used to interpret the effect size of these univariate correlation coefficients as small (r = .10), medium (r = .30), and large (r = .50).

To further examine the construct validity for the HSI and its factors, a multivariate analysis of covariance, adjusted for age and race/ethnicity, examined mean differences in HSI-S, HSI-I, and HSI-T scores between the 3 diagnostic groups (ie, CDH, CID, and OSD) and a multivariate analysis of variance examined mean

differences in HSI-S, HSI-I, and HSI-T scores between 3 CDH subgroups (ie, psychiatric hypersomnia, idiopathic hypersomnia, and narcolepsy).

Finally, receiver operating characteristic analyses with calculation of the area under the curve (AUC) were used to examine the criterion and concurrent validity of the HSI and its factors for the presence of CDH (vs ESS score < 10; vs CID; vs OSD) as criterion variables. Sensitivity and specificity values were extracted for HSI-S, HSI-I, and HSI-T scores against the criterion variables to inform potential cutoff scores. A cutoff score was considered optimal based on its best balance between sensitivity and specificity (ie, it simultaneously maximized both indices).

RESULTS

Characteristics of the patient sample

The demographic characteristics of the patient sample are presented in **Table 1**. Participants were predominantly middle-aged (median age 44 years), female (70.9%), and non-Hispanic White (80.4%). The average body mass index was approximately 30 kg/m^2, with 29.7% of participants classified as normal weight, 25.9% as overweight, and 44.3% as obese. Approximately 24.6% of the patients had received a diagnosis of CDH, 46.2% CID, and 29.1% OSD. Among the latter, the most common diagnoses were sleep-related breathing disorders (58.7%) followed by circadian rhythm sleep-wake disorders (30.4%) and sleep-related movement disorders (28.3%), which could be comorbid between each other (eg, sleep-related breathing and movement disorders). There were significant differences across the CDH, CID, and OSD groups in terms of age and race/ethnicity distribution. Specifically, participants with CDH (37.72 (15.01)) were significantly younger than those with CID (44.16 (16.22), $P = .04$) or OSD (49.46 (16.13), $P = .01$). Furthermore, those with CID ($P = .03$), and to lesser extent those with CDH ($P = .07$), were more likely to belong to a racial/ethnic minority compared to those with OSD. These demographic variables were adjusted for when we compared these 3 diagnostic groups between each other on the study outcomes.

Reliability and factorial validity

The HSI showed satisfactory internal consistency (Cronbach's $\alpha = 0.79$) and item-total correlations ($r = .42$–$.67$), except for item 1a (sleeping too much at night; $r = .17$). The results of the exploratory factor analysis are also presented in **Table 2**, which indicated that 2 factors were retained. These 2 factors accounted for 56.2% of the total variance. The first factor contained all of the items associated with distress or impairment (HSI-I), with item factor loadings ranging from 0.71–0.83. The second factor contained items related to essential hypersomnia symptoms (HSI-S), with item factor loadings ranging from 0.56–0.77. The internal consistency of the HSI-I factor was high ($\alpha = 0.81$) and lower for the HSI-S factor ($\alpha = 0.69$) in the overall patient sample. However, when stratified by the presence of each disorder, the internal consistency of the HSI-S factor was higher among patients

Table 1—Demographic and clinical characteristics of the patient sample.

	All (n = 158)	CID (n = 73)	OSD (n = 46)	CDH (n = 39)	*P*
Age, y	44.11 (16.38)	44.16 (16.22)	49.46 (16.13)	37.72 (15.01)	.004
≤ 34	29.1%	30.1%	15.2%	43.6%	.007
35–44	22.8%	20.5%	17.4%	33.3%	—
45–59	25.9%	26.0%	34.8%	15.4%	—
≥ 60	22.2%	23.3%	32.6%	7.7%	—
Sex					
Male	29.1%	32.9%	30.4%	20.5%	.380
Female	70.9%	67.1%	69.6%	79.5%	—
Race/ethnicity					
Non-Hispanic White	80.4%	75.3%	91.3%	76.9%	.084
Racial/ethnic minority	19.6%	24.7%	8.7%	23.1%	
BMI, kg/m^2	30.34 (7.89)	29.26 (6.89)	31.66 (8.96)	30.80 (8.20)	.251
Normal weight	29.7%	31.5%	28.3%	28.2%	.470
Overweight	25.9%	31.5%	19.6%	23.1%	—
Obese	44.3%	37.0%	52.2%	48.7%	—
SBP, mm Hg	121.28 (14.93)	120.21 (14.19)	124.52 (16.47)	119.49 (14.15)	.212
DBP, mm Hg	72.37 (11.13)	71.32 (9.74)	71.83 (11.42)	75.00 (12.96)	.231

Data are mean (standard deviation) for continuous variables and percentage for categorical variables. *P* value for differences between groups, from Pearson chi-square test for categorical variables and from analysis of variance for continuous variable. BMI = body mass index, CDH = central disorders of hypersomnolence, CID = chronic insomnia disorder, DBP = diastolic blood pressure, OSD = other sleep disorder, SBP = systolic blood pressure.

Table 2—Reliability and factorial validity of the HSI.

Item	Cronbach's α If Item Removed	Item—Total Correlation	Factor I (HSI-I)	Factor II (HSI-S)
1a—sleep too much at night	0.807	0.168	—	0.557
1b—difficulty waking up	0.771	0.487	—	0.691
1c—sleep during the day	0.774	0.451	—	0.773
1d—sleepy during the daytime	0.756	0.573	—	0.579
2—satisfied/dissatisfied	0.781	0.398	0.733	—
3—interference	0.743	0.672	0.832	—
4—noticeability	0.740	0.670	0.709	—
5—worried/distressed	0.777	0.423	0.815	—
6—sleep attacks	0.771	0.475	—	0.594
Cronbach's α	0.790	—	0.814	0.691
Explained variance	—	—	30.4%	25.8%
	Total (HSI-T)	—	**Factor I (HSI-I)**	**Factor II (HSI-S)**
Mean (SD)	18.15 (6.57)	—	11.20 (3.64)	6.94 (4.14)
Q1	14.00	—	9.00	3.75
Q2	18.00	—	11.50	7.00
Q3	23.00	—	15.00	9.00

Loadings in factor I or factor II below 0.40 are not shown. HSI = Hypersomnia Severity Index, HSI-I = Hypersomnia Severity Index-impairment, HSI-S = Hypersomnia Severity Index-symptoms, HSI-T = Hypersomnia Severity Index-total, Q1 = 25th percentile, Q2 = 50th percentile, Q3 = 75th percentile, SD = standard deviation.

with CDH (α = 0.70), whereas it was lower among patients with CID (α = 0.59) or OSD (α = 0.57).

Construct validity: patient-reported outcomes

Table 3 presents convergent validity results for the HSI and its factors with related measures of daytime sleepiness, fatigue, and depression. HSI-S scores showed large, statistically significant correlations with ESS scores and medium, statistically significant correlations with FFS and DASS-D scores. In contrast, the HSI-I scores showed large, statistically significant correlations with FFS scores and medium, statistically significant correlations with DASS-D and ESS scores.

Evidence for divergent validity for the HSI and its factors with unrelated measures of sleep reactivity, sleep effort, and sleep-incompatible behaviors (SHPS-I) was also examined and is presented in Table 3. HSI-S scores showed nonsignificant correlations with FIRST, GSES, or SHPS-I scores. In contrast, HSI-I scores showed small-to-medium, statistically significant correlations with GSES, FIRST, and SHPS-I scores.

Construct validity: diagnostic groups

Further construct validity across diagnostic groups (Table 4) showed significantly (P = .004) higher HSI-T scores in patients with CDH and marginally (P = .066) lower scores in patients with CID as compared to those with OSD. These divergent scores were differentially driven by the contribution of HSI-S vs HSI-I scores. Specifically, patients with CDH showed significantly higher HSI-S scores as compared to patients with CID (P < .0001) and

Table 3—Construct validity of the HSI: patient-reported outcomes.

	Mean (SD)	HSI-S	HSI-I	HSI-T
ESS	8.30 (5.18)	0.692**	0.387**	0.651**
FFS	17.83 (7.44)	0.441**	0.690**	0.664**
DASS-D	11.91 (10.94)	0.329**	0.395**	0.426**
FIRST	23.27 (6.67)	0.152	0.391**	0.314**
GSES	7.81 (3.59)	0.023	0.491**	0.289**
SHPS-I	11.03 (3.92)	0.130	0.249**	0.220**

Data are Pearson correlation coefficients, unless otherwise stated. **$P \leq .01$. DASS-D = Depression Anxiety and Stress Scale-depression, ESS = Epworth Sleepiness Scale, FFS = Flinders Fatigue Scale, FIRST = Ford Insomnia Response to Stress Test, GSES = Glasgow Sleep Effort Scale, HSI = Hypersomnia Severity Index, HSI-I = Hypersomnia Severity Index-impairment, HSI-S = Hypersomnia Severity Index-symptoms, HSI-T = Hypersomnia Severity Index-total, SD = standard deviation, SHPS-I = Sleep Hygiene Practices Scale-incompatible behaviors.

Table 4—Construct validity of the HSI: diagnostic groups.

	CID[a] (n = 73)	OSD[a] (n = 46)	CDH[a] (n = 39)	Psychiatric[b] (n = 14)	Idiopathic[b] (n = 15)	Narcolepsy[b] (n = 10)
HSI-S	5.66 (3.67)	6.69 (3.80)	9.65 (3.75)	8.79 (3.31)	8.87 (3.38)	13.40 (3.27)
HSI-I	10.35 (3.59)	11.46 (3.66)	12.50 (3.62)	11.71 (2.43)	12.40 (4.82)	13.90 (2.60)
HSI-T	16.01 (6.06)	18.14 (6.17)	22.15 (6.12)	20.50 (5.11)	21.27 (7.24)	27.30 (4.99)

[a]Data are means (standard deviation) adjusted for race/ethnicity and age (see text). [b]Data are means (standard deviation). CDH = central disorders of hypersomnolence, CID = chronic insomnia disorder, ESS = Epworth Sleepiness Scale, HSI = Hypersomnia Severity Index, HSI-I = Hypersomnia Severity Index-impairment, HSI-S = Hypersomnia Severity Index-symptoms, HSI-T = Hypersomnia Severity Index-total, OSD = other sleep disorder.

with OSD ($P = .001$), whereas patients with CID and OSD did not significantly differ between each other after adjusting for covariates ($P = .150$). In contrast, patients with CDH showed significantly higher HSI-I scores as compared to patients with CID ($P = .003$) but not as compared to patients with OSD ($P = .200$). Patients with CID and OSD did not significantly differ between each other in HSI-I scores ($P = .107$). Moreover, as shown in **Table 4**, patients with narcolepsy showed significantly higher HSI-T scores compared to those with psychiatric ($P = .010$) or idiopathic hypersomnia ($P = .019$), whereas patients with psychiatric vs idiopathic hypersomnia did not significantly differ between each other ($P = .733$). Patients with narcolepsy showed significantly higher HSI-S scores as compared to patients with psychiatric ($P = .002$) or idiopathic hypersomnia ($P = .002$), whereas these latter 2 groups did not significantly differ between each other ($P = .948$). In contrast, HSI-I scores did not significantly differ either between those with narcolepsy compared to idiopathic ($P = .313$) or psychiatric hypersomnia ($P = .150$) or between the latter 2 groups ($P = .610$).

Criterion and concurrent validity

Receiver operating characteristic analyses were conducted to determine the concurrent validity of the HSI and its factors in identifying the presence of CDH. As shown in **Table 5**, the HSI-S (AUC, 0.85; 95% confidence interval [CI], 0.76–0.92) and HSI-T (AUC, 0.83; 95% CI, 0.75–0.92) performed better than the HSI-I (AUC, 0.71; 95% CI, 0.59–0.78) when identifying the presence of CDH vs an ESS score < 10. Similar findings were observed in the comparison group (the CID group), with the HSI-S (AUC, 0.80; 95% CI, 0.71–0.88) and HSI-T (AUC, 0.79; 95% CI, 0.70–0.89) performing better than the HSI-I (AUC, 0.69; 95% CI, 0.59–0.78) when identifying the presence of CDH vs CID. In contrast, the HSI-S (AUC, 0.76; 95% CI, 0.65–0.86), HSI-T (AUC, 0.72; 95% CI, 0.60–0.83), and HSI-I (AUC, 0.61; 95% CI, 0.49–0.73) all performed slightly worse when identifying the presence of CDH vs OSD; notably, 35% of patients with OSD had an ESS score > 10. Together, these data showed that an HSI-S score of 8 points had the best balance in sensitivity (0.82) and specificity (0.78) to identify CDH in this clinical patient sample (AUC, 0.85).

DISCUSSION

The present study aimed to examine the reliability, factor structure, construct, and criterion validity of the HSI, a patient-reported measure of hypersomnolence, in a clinical sample of patients with sleep disorders. Our results support that the HSI assesses 2 factors, hypersomnia symptoms (HSI-S) and associated distress and impairment (HSI-I), with optimal reliability, as previously reported.[10] Its construct and criterion validity are also supported by its convergent and divergent association with other patient-reported outcomes, by differences between patients diagnosed with CDH vs CID or OSD, and by differences between patients diagnosed with narcolepsy vs idiopathic and psychiatric hypersomnia. Our results also provide guiding cutoff scores for the HSI and its factors when implemented as part of the routine assessment of clinically diverse sleep disorders. Together, these findings support the HSI as a robust measure of hypersomnia and suggest the need for studies establishing severity thresholds and its sensitivity to treatment.

Analyses supported a 2-factor structure of the HSI, accounting for approximately 56% of the total variance, which is strikingly similar to the results found in a sample of undergraduate

Table 5—Criterion validity of the HSI.

	CDH vs ESS Score < 10	CDH vs CID	CDH vs OSD
HSI-S	0.85 (7.5), 0.82/0.78	0.80 (8.5), 0.69/0.78	0.76 (8.5), 0.69/0.72
HSI-I	0.71 (11.5), 0.59/0.62	0.69 (11.5), 0.59/0.62	0.61 (12.5), 0.54/0.52
HSI-T	0.83 (19.5), 0.77/0.77	0.79 (19.5), 0.77/0.70	0.72 (20.5), 0.64/0.63

Data are area under the curve (cutoff score) with best balance between sensitivity/specificity. CDH = central disorders of hypersomnolence, CID = chronic insomnia disorder, ESS = Epworth Sleepiness Scale, HSI = Hypersomnia Severity Index, HSI-I = Hypersomnia Severity Index-impairment, HSI-S = Hypersomnia Severity Index-symptoms, HSI-T = Hypersomnia Severity Index-total, OSD = other sleep disorder.

students.[10] Internal consistency of the HSI and of the HSI-I factor was high and slightly lower for the HSI-S factor, also consistent with previous research.[10] This finding could potentially be explained by the heterogeneity of the patients in the clinically diverse sleep disorders sample and by the variety of symptoms assessed within the HSI-S factor. Indeed, the internal reliability of the HSI-S factor was inadequate among those with CID and OSD and optimal among those with CDH, which supports its use as a severity measure and potential screening instrument for CDH.

Convergent validity for the HSI was good, and particularly for the HSI-S factor. HSI-S scores correlated positively and strongly with measures of sleepiness, fatigue, and depressive symptoms. Notably, HSI-S scores did not significantly correlate with measures of sleep reactivity, sleep effort, or sleep-incompatible behaviors, all measures of stress, arousal, and/or behavioral factors inducing sleep disturbance that should be unrelated to central hypersomnolence, particularly the excessive need to sleep. In contrast, HSI-I scores correlated positively and strongly not only with measures of sleepiness, fatigue, and depressive symptoms but also with sleep reactivity, sleep effort, and sleep-incompatible behaviors. These results indicate adequate convergent and divergent validity for the HSI-S factor and adequate convergent validity and poor divergent validity for the HSI-I factor. The stronger divergent validity of the HSI-S factor was further supported by the significantly higher scores in patients with CDH as compared to those with either CID or OSD, whereas HSI-I scores were similar between patients with CDH and those with OSD. In addition, the divergent validity of the HSI-S factor was further supported by the significantly higher scores in patients with narcolepsy as compared to those with either psychiatric or idiopathic hypersomnia, whereas HSI-I scores were similar across these 3 CDH subgroups.

These data are important from a clinical implementation standpoint. It seems that when the HSI is used as a screening or global assessment measure, many individuals with sleep disorders will score high on items 2–5 (HSI-I factors: satisfaction/dissatisfaction with sleep pattern, interference with daily functioning, how noticeable the sleep problem is to others, and worry/distress about sleep problem) because they may not necessarily limit those items to the hypersomnolence symptoms covered in items 1a through 1d and 6 (sleeping too much at night, feeling sleepy during the daytime, and having "sleep attacks"/unintended sleep in inappropriate situations) in their self-reporting, suggesting that the interpretation of the HSI-T score may require the involvement of the clinician to assure the presence of clinically significant hypersomnia symptoms. Nevertheless, this issue does not preclude the use of the HSI as a screening device in the general population or in clinical patient samples.

In the present study, guiding cutoff scores for the HSI and its factors were also delineated. Our results indicated that the HSI-S score, compared to the HSI-T and HSI-I score, showed optimal diagnostic confidence, as determined by the intersection point of sensitivity and specificity, when identifying individuals diagnosed with CDH as compared to those without excessive daytime sleepiness (ESS score < 10) or those diagnosed with CID (regardless of their ESS score). Furthermore, it is evident that the predictive value of the HSI-T score in the current patient sample was mostly driven by HSI-S scores, as indicated by their AUC of 0.85 and a much lower AUC of 0.71 for HSI-I scores. We propose that an HSI-S score ≥ 8 points is most optimal to screen individuals for clinically significant hypersomnia symptoms and potential CDH when the HSI is administered as part of the routine assessment of patients with different sleep disorders that may or may not present with excessive daytime sleepiness. In such clinical use, it seems that the HSI-S score could be interpreted first and should inform the interpretation of the HSI-I and HSI-T scores as per overall severity purposes. We propose that this 2-step process will make the HSI a unique, clinically useful measure allowing clinicians to establish first the presence of the essential symptoms (HSI-S) and, thereafter, the severity of hypersomnia based on its associated impact (HSI-I). However, there is still a need to derive such severity cutoff scores by including individuals without sleep disorders and otherwise good sleepers in future studies.

The results of the current study should be interpreted in light of some limitations. The cross-sectional nature of the study prevents any definitive conclusions about the direction of the relationship between hypersomnia and clinically relevant constructs (concurrent validity); therefore, future longitudinal research is needed to examine its predictive validity. In addition, time-stamped data on medication status at the time of assessment were not obtained for the current patient sample, and it is possible that some patients may have already been on sleep and/or alertness medications. Therefore, it is important for future research to examine the performance of the HSI in medicated and unmedicated patients with CDH and its sensitivity to the effects of alertness medications. Although this sample of patients with sleep disorders adds to previous data restricted to patients with psychiatric disorders, there is still a need to replicate these findings in population-based cohorts and clinical patient samples that include good sleeping control patients with similar demographic characteristics. Furthermore, the sample size of each subtype of patients with CDH was limited, and further studies in individuals with different types of narcolepsy and hypersomnia are still needed. Given that this was a retrospective chart review and given the known problems of extracting diagnoses from electronic medical records, testing the reliability of the diagnoses (ie, interrater or against external criteria) was not possible; however, as previously mentioned, sleep/circadian diagnoses were extracted when they were made by board-certified sleep physicians or sleep psychologists. Finally, findings should be taken cautiously given the specific demographic characteristics of the patient sample, which may not be representative of other clinical patient samples or the general population at large; based on the current and previous study,[10] there is a need to test the performance of the HSI in racially/ethnically diverse patient samples, particularly including Black/African American patients and Hispanic/Latino patients.

In summary, the HSI is a brief measure of hypersomnia with satisfactory indices of reliability and validity that can be used in clinical samples of patients with sleep disorders. Its construct and criterion validity are supported by its convergent and divergent association with other patient-reported outcomes and the adequate sensitivity/specificity of its HSI-S factor to reliably identify individuals with CDH. The guiding cutoff score of 8 for the HSI-S

can aid clinicians in screening efforts, case formulation, and targeted treatments, particularly when other measures used in routine clinical practice (eg, the ESS) do not provide sufficient information to establish the presence of hypersomnolence. Future studies are needed to replicate the psychometric properties of the HSI and the cutoffs determined herein in representative population-based patient samples and to develop cutoff scores informing severity thresholds and estimating the sensitivity of the HSI to pharmacological and behavioral treatments.

ABBREVIATIONS

AUC, area under the curve
CDH, central disorders of hypersomnolence
CI, confidence interval
CID, chronic insomnia disorder
DASS-D, Depression Anxiety Stress Scale-depression
ESS, Epworth Sleepiness Scale
FFS, Flinders Fatigue Scale
FIRST, Ford Insomnia Response to Stress Test
GSES, Glasgow Sleep Effort Scale
HSI, Hypersomnia Severity Index
HSI-I, Hypersomnia Severity Index-impairment
HSI-S, Hypersomnia Severity Index-symptoms
HSI-T, Hypersomnia Severity Index-total
OSD, other sleep disorder
SHPS-I, Sleep Hygiene Practice Scale, sleep-incompatible behaviors

REFERENCES

1. Ohayon MM. From wakefulness to excessive sleepiness: what we know and still need to know. *Sleep Med Rev.* 2008;12(2):129–141.
2. Slater G, Steier J. Excessive daytime sleepiness in sleep disorders. *J Thorac Dis.* 2012;4(6):608–616.
3. Ohayon MM. Determining the level of sleepiness in the American population and its correlates. *J Psychiatr Res.* 2012;46(4):422–427.
4. Bixler EO, Vgontzas AN, Lin HM, Calhoun SL, Vela-Bueno A, Kales A. Excessive daytime sleepiness in a general population sample: the role of sleep apnea, age, obesity, diabetes, and depression. *J Clin Endocrinol Metab.* 2005;90(8):4510–4515.
5. Fernandez-Mendoza J, Vgontzas AN, Kritikou I, Calhoun SL, Liao D, Bixler EO. Natural history of excessive daytime sleepiness: role of obesity, weight loss, depression, and sleep propensity. *Sleep*. 2015;38(3):351–360.
6. Drake C, Roehrs T, Breslau N, et al. The 10-year risk of verified motor vehicle crashes in relation to physiologic sleepiness. *Sleep*. 2010;33(6):745–752.
7. Ohayon MM. Determining the level of sleepiness in the American population and its correlates. *J Psychiatr Res.* 2012;46(4):422–427.
8. Stevenson MR, Elkington J, Sharwood L, et al. The role of sleepiness, sleep disorders, and the work environment on heavy-vehicle crashes in 2 Australian states. *Am J Epidemiol.* 2014;179(5):594–601.
9. Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep*. 1991;14(6):540–545.
10. Kaplan KA, Plante DT, Cook JD, Harvey AG. Development and validation of the Hypersomnia Severity Index (HSI): a measure to assess hypersomnia severity and impairment in psychiatric disorders. *Psychiatry Res.* 2019;281:112547.
11. Dauvilliers Y, Evangelista E, Barateau L, et al. Measurement of symptoms in idiopathic hypersomnia: the Idiopathic Hypersomnia Severity Scale. *Neurology.* 2019;92(15):e1754–e1762.
12. American Academy of Sleep Medicine. *International Classification of Sleep Disorders.* 3rd ed. Darien, IL: American Academy of Sleep Medicine; 2014.
13. Gradisar M, Lack L, Richards H, et al. The Flinders Fatigue Scale: preliminary psychometric properties and clinical sensitivity of a new scale for measuring daytime fatigue associated with insomnia. *J Clin Sleep Med.* 2007;3(7):722–728.
14. Lovibond SH, Lovibond PF. *Manual for the Depression Anxiety Stress Scales.* 2nd ed. Sydney, Australia: Psychology Foundation of Australia; 1995.
15. Antony MM, Bieling PJ, Cox BJ, Enns MW, Swinson RP. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychol Assess.* 1998;10(2):176–181.
16. Drake C, Richardson G, Roehrs T, Scofield H, Roth T. Vulnerability to stress-related sleep disturbance and hyperarousal. *Sleep.* 2004;27(2):285–291.
17. Kalmbach DA, Pillai V, Arnedt JT, Drake CL. Identifying at-risk individuals for insomnia using the Ford Insomnia Response to Stress Test. *Sleep.* 2016;39(2):449–456.
18. Broomfield NM, Espie CA. Towards a valid, reliable measure of sleep effort. *J Sleep Res.* 2005;14(4):401–407.
19. Hertenstein E, Nissen C, Riemann D, Feige B, Baglioni C, Spiegelhalder K. The exploratory power of sleep effort, dysfunctional beliefs and arousal for insomnia severity and polysomnography-determined sleep. *J Sleep Res.* 2015;24(4):399–406.
20. Lin SC, Cheng CP, Yang CM, Hsu SC. Psychometric properties of the Sleep Hygiene Practice Scale. *Sleep.* 2007;30:A262.
21. Yang CM, Lin SC, Hsu SC, Cheng CP. Maladaptive sleep hygiene practices in good sleepers and patients with insomnia. *J Health Psychol.* 2010;15(1):147–155.
22. Brown TA. *Confirmatory Factor Analysis for Applied Research.* 2nd ed. New York, NY: The Guilford Press; 2015.

ACKNOWLEDGMENTS

The authors thank the sleep technicians and staff of the Sleep Research and Treatment Center for their effort and support with this project.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication January 14, 2021
Submitted in final revised form May 7, 2021
Accepted for publication May 7, 2021
Address correspondence to: Julio Fernandez-Mendoza, PhD, DBSM, Sleep Research and Treatment Center, Department of Psychiatry and Behavioral Health, Penn State Health Milton S. Hershey Medical Center, Penn State University College of Medicine, 500 University Drive H073, Hershey, PA 17603; Email: jfmendoza@psu.edu

DISCLOSURE STATEMENT

All authors have seen and approved the final manuscript. Work for this study was performed at the Sleep Research and Treatment Center of Penn State Health Milton S. Hershey Medical Center, Hershey, PA. Dr. Fernandez-Mendoza has received unrelated research support from the National Heart, Lung, and Blood Institute (R01HL136587), the National Institute of Mental Health (R01MH118308), the American Heart Association (14SDG19830018), the Patient-Centered Outcome Research Institute (subcontract with grant number CER2018C213262), and Pfizer Central Research (NCT03665454). Dr. Plante has received unrelated research support from the American Sleep Medicine Foundation, the National Institutes of Health, the University of Illinois at Chicago Occupational and Environmental Health and Safety Education and Research Center (funded by the National Institute for Occupational Safety and Health), and the Madison Educational Partnership. Dr. Plante has served as a consultant for Teva Pharmaceuticals Australia, a consultant for Harmony Biosciences, and as a consultant/medical advisory board member for Jazz Pharmaceuticals. All other authors report no conflicts of interest.