ONLINE FIRST

Testing the Reliability and Validity of *DSM-IV-TR* and *ICSD-2* Insomnia Diagnoses

Results of a Multitrait-Multimethod Analysis

Jack D. Edinger, PhD; James K. Wyatt, PhD; Edward J. Stepanski, PhD; Maren K. Olsen, PhD; Karen M. Stechuchak, MS; Colleen E. Carney, PhD; Ambrose Chiang, MD; M. Isabel Crisostomo, MD; Margaret D. Lineberger, PhD; Melanie K. Means, PhD; Rodney A. Radtke, MD; William K. Wohlgemuth, PhD; Andrew D. Krystal, MD

Context: Distinctive diagnostic classification schemes for insomnia diagnoses are available, but the optimal insomnia nosology has yet to be determined.

Objectives: To test the reliability and validity of insomnia diagnoses listed in the American Psychiatric Association's *DSM-IV-TR* and the *International Classification of Sleep Disorders*, second edition (*ICSD-2*).

Design: Multitrait-multimethod correlation design.

Setting: Two collaborating university medical centers, with recruitment from January 2004 to February 2009.

Participants: A total of 352 adult volunteers (235 of whom were women) who met research diagnostic criteria for insomnia disorder.

Main Outcome Measures: Goodness-of-fit ratings of 10 *DSM-IV-TR* and 37 *ICSD-2* insomnia diagnoses for each patient. Ratings were provided by 3 clinician pairs who used distinctive assessment methods to derive diagnostic impressions. Correlations computed within and across clinician pairs were used to test reliability and validity of diagnoses.

Results: Findings suggested that the best-supported *DSM-IV-TR* insomnia categories were insomnia related to another mental disorder, insomnia due to a general medical condition, breathing-related sleep disorder, and circadian rhythm sleep disorder. The category of primary insomnia appeared to have marginal reliability and validity. The best-supported *ICSD-2* categories were the insomnias due to a mental disorder and due to a medical condition, obstructive sleep apnea, restless legs syndrome, idiopathic insomnia, and circadian rhythm sleep disorder–delayed sleep phase type. Psychophysiological insomnia and inadequate sleep hygiene received much more variable support across sites, whereas the diagnosis of paradoxical insomnia was poorly supported.

Conclusions: Both the *DSM-IV-TR* and *ICSD-2* provide viable insomnia diagnoses, but findings support selected subtypes from each of the 2 nosologies. Nonetheless, findings regarding the frequently used *DSM-IV-TR* diagnosis of primary insomnia and its related *ICSD-2* subtypes suggest that their poor reliability and validity are perhaps due to significant overlap with comorbid insomnia subtypes. Therefore, alternate diagnostic paradigms should be considered for insomnia classification.

Arch Gen Psychiatry. 2011;68(10):992-1002.
Published online June 6, 2011.
doi:10.1001/archgenpsychiatry.2011.64

Author Affiliations are listed at the end of this article.

THERE ARE SEVERAL DIAGNOSTIC nosologies for insomnia[1-6] designed to systematize descriptions of patients, facilitate communication among health care practitioners, guide treatment choices, predict clinical course, and standardize research.[7] These nosologies differ markedly in their complexity and reliance on information external to the clinical interview. The American Psychiatric Association's *DSM* (*DSM-III-R*, *DSM-IV*, and *DSM-IV-TR*)[3,4] describes a few global insomnia diagnoses and relies primarily on clinical interview. In contrast, the *International Classification of Sleep Disorders* (*ICSD*) and the *International Classification of Sleep Disorders*, second edition (*ICSD-2*)[5,8] delineate numerous primary and secondary insomnia subtypes and incorporate findings from interview and laboratory tests.

See also page 990

These divergent classification schemes are products of discrepant views about how many subtypes are needed to describe all individuals with insomnia. Proponents of the *DSM-IV-TR* nosology argue that many *ICSD-2* insomnia subtypes have little empirical substantiation and should be subsumed within broader categories.[9-11] However, the *DSM-IV-TR* system allows for considerable heterogeneity within

diagnostic categories and may not provide optimal discrimination among distinctive insomnia disorders.[12-18] What is clear is that the 2 nosologies result in markedly discordant classifications when applied to the same sample of patients with insomnia.[19] This state of affairs creates costly variability in assessment and management of patients with insomnia and needless disunity within the insomnia research literature.

Whether the *DSM-IV-TR* or *ICSD-2* offers a more accurate scheme for insomnia classification and diagnosis remains unknown. The few studies[20-24] that assessed reliability of *DSM-IV-TR* or *ICSD-2* diagnostic categories have found only modest reliability for the insomnia subtypes they evaluated. Reliability data are unavailable for many of the diagnoses in each system. Moreover, available literature[21,25-27] provides indirect and limited support for a selected subset of *DSM-IV-TR* and *ICSD-2* diagnoses. For example, one study[26] showed that treatment recommendations of clinicians vary as a function of the *DSM-IV* and *ICSD* diagnoses they assign. Other studies[21,25] comparing patient groupings resulting from standard clinical classification with the groupings resulting from statistical clustering procedures have shown some (albeit minimal) congruence between these 2 classification approaches. Whereas such studies represent proxies for testing the validity of insomnia diagnoses, formal omnibus empirical tests of these nosologies have not yet been conducted to our knowledge.

Nonetheless, the *DSM-IV-TR* and *ICSD-2* have continued to enjoy widespread clinical and research use. Given the existence of these 2 discordant nosologies, insomnia diagnosis and treatment remain a hit-and-miss process guided more by clinicians' instincts and beliefs about various insomnia subtypes than by a well-validated diagnostic system. Clearly, research to ascertain the most viable insomnia nosology is sorely needed. This dual-site study was conducted with the following aims: (1) to determine and compare the reliabilities of *DSM-IV-TR* and *ICSD-2* insomnia subtypes; and (2) to derive and compare convergent validity (CV) and discriminant validity (DV) indices for the *DSM-IV-TR* and *ICSD-2* insomnia diagnoses. Our overarching goal was to ascertain the optimal scheme for insomnia classification.

METHODS

DESIGN

This study was conducted at Duke University Medical Center and Rush University Medical Center using a multitrait-multimethod[28] design. Multiple insomnia diagnoses were assessed in a research cohort using multiple assessment methods. Within each study site, 6 sleep specialists were grouped to form 3 clinician pairs. Each pair was then assigned 1 of 3 assessment approaches to use throughout the study for discerning the insomnia diagnosis (or diagnoses) of each participant they interviewed. One pair used solely a structured sleep interview for discerning diagnoses; the second pair used standard unstructured clinical interviews; and the third pair relied on unstructured clinical interviews combined with polysomnographic (PSG) data. The latter 2 pairs were also given access to information from sleep diaries and sleep history questionnaires completed by study participants. Each clinician formulated impressions independently without knowledge of the other clinicians' impressions. The multiple traits considered were 10 *DSM-IV-TR* and 37 *ICSD-2* insomnia diagnoses, which represent all *DSM-IV-TR* and *ICSD-2* insomnia diagnoses that can be ascertained via interview. Following their interviews, clinicians rated how well each of these 47 insomnia diagnoses fit each patient. A standard multitrait-multimethod[28] correlational analysis was then applied to these ratings to test the reliability and validity of the insomnia diagnoses considered. The institutional review boards of the collaborating medical centers reviewed and approved the study protocol. Participants provided written informed consent and received parking expenses plus a maximum $400 payment for participation.

Table 1. Demographic Characteristics of Study Sample

Characteristic	Overall (N=352)	Duke University Medical Center (n=201)	Rush University Medical Center (n=151)
Age, mean (SD), y	46.44 (14.40)	49.25 (14.83)	42.70 (12.95)
Sex, No. (%)			
Male	117 (33.24)	67 (33.33)	50 (33.11)
Female	235 (66.76)	134 (66.67)	101 (66.89)
Race, No. (%)			
White	212 (60.23)	140 (69.65)	72 (47.68)
African American	112 (31.82)	56 (27.86)	56 (37.09)
American Indian or Alaskan	1 (0.28)	1 (0.50)	0
Asian or Pacific Islander	13 (3.69)	2 (1.00)	11 (7.28)
Other	14 (3.98)	2 (1.00)	12 (7.95)
Education, mean (SD), y[a]	15.08 (3.10)	15.54 (2.77)	14.47 (3.41)
Marital status, No. (%)[a]			
Single	135 (38.46)	53 (26.37)	82 (54.67)
Married or living as married	142 (40.46)	99 (49.25)	43 (28.67)
Divorced, separated, or widowed	74 (21.08)	49 (24.38)	25 (16.67)

[a] One subject is missing education and 1 subject has missing data for marital status. Percentage calculations exclude subjects with missing data.

PARTICIPANTS

Recruitment occurred between January 2004 and February 2009 through posted announcements and physician referrals. Included individuals (1) met Research Diagnostic Criteria[26] for insomnia disorder, (2) were aged 18 years or younger, and (3) spoke English fluently. Excluded individuals (1) had an unstable or life-threatening medical condition, (2) were imminently suicidal, (3) scored 24 or lower on the Mini-Mental State Examination, or (4) were previously evaluated by any study clinicians. Of the 425 individuals enrolled, 8 were removed because they did not meet study selection criteria, 50 withdrew before beginning any study interviews, and 15 failed to complete all interviews. The remaining 352 participants (201 from Duke University Medical Center and 151 from Rush University Medical Center) composed the final sample. **Table 1** shows demographic characteristics for the sample.

POLYSOMNOGRAPHY

Participants underwent 2 consecutive nights of PSG with a monitoring montage consisting of 2 channels of electroencephalography (C3-M2, Oz-Cz), 1 chin electromyography channel, 2

channels of electro-oculography (left eye to M1, right eye to M2), 1 channel of airflow (nasal-oral thermistor), 2 channels of respiratory effort (thoracic and abdominal impedance), 1 channel of pulse oximetry, 2 channels of anterior tibialis electromyography (right and left legs), and 1 channel of body position monitoring. Participants followed their customary bedtimes and rising times on PSG nights. Those who occasionally used hypnotics underwent PSG sessions without such medications, whereas those who used hypnotics 3 or more nights per week or were taking antidepressants and/or anxiolytics underwent PSG sessions with these medications.

All PSG sessions were scored using traditional scoring criteria for sleep stages, apneas and hypopneas, periodic limb movements, and related arousals.[29-31] Summary data including respiratory parameters (apnea-hypopnea index, desaturation index, etc), periodic limb movement indices (number of movements with and without arousals), and types and dosages of medications taken on PSG nights were included in a report made available to 1 clinician pair at each study site.

ELECTRONIC SLEEP DIARY

Participants recorded sleep data for 2 weeks using a handheld computer. Those who had difficulty using this device completed paper sleep diaries. The computer presented questions about each night's bedtime, sleep onset latency, number and length of nocturnal awakenings, time of final awaking, rising time, and sleep medication and alcohol use. Also, respondents' ratings (10-point scale) of sleep quality and how rested they felt on arising were acquired. Diary data were downloaded (or hand-entered for paper diaries) into a PC computer, and a printout of daily values and 2-week averages was generated showing the following: bedtime, sleep onset latency, number of nocturnal awakenings, time awake after sleep onset and prior to final awakening, time of final awakening, time of rising out of bed, total sleep time, total time awake, time in bed, sleep efficiency (total sleep time/time in bed $\times$ 100), sleep quality, restedness on arising, medication and alcohol use, and the time when diary entries were made.

SLEEP HISTORY QUESTIONNAIRE

Participants completed a 10-page questionnaire. This solicited information about their demographic characteristics, sleep complaints, medical and psychiatric history, and treatment history.

STRUCTURED SLEEP INTERVIEW

The Duke Structured Interview for Sleep Disorders (DSISD) was used by 1 clinician pair at each site to derive participants' insomnia diagnoses. The DSISD incorporates criteria for ascertaining *DSM-IV-TR* and *ICSD-2* sleep disorder diagnoses and is divided into 4 modules: insomnia-related disorders, excessive daytime sleepiness–related disorders, sleep/wake schedule disorders, and parasomnias. The insomnia module assesses insomnias related to other mental disorders, general medical disorders, substance abuse, circadian rhythm disorders, restless legs syndrome, inadequate sleep hygiene, etc. Screening questions allow sections within a module to be skipped contingent on definitive negative answers from a respondent. However, interviewers may continue sections when answers to screening questions are ambiguous. Previous studies show that the DSISD has acceptable reliability and validity for *DSM-IV-TR* and *ICSD-2* insomnia diagnoses.[32,33]

ELECTRONIC DIAGNOSTIC RATING FORMS

The rating forms consisted of the series of 10 *DSM-IV-TR* and 37 *ICSD-2* diagnoses presented on the screen of a specially programmed handheld computer. Each diagnosis appeared individually accompanied by a 100-pixel visual analog scale (VAS) labeled "doesn't fit at all" at its left extreme and "fits extremely well" at its right extreme. Clinicians considered each diagnosis separately and decided how well it fit the participant in question. Clinicians moved a pointer on the VAS to indicate the goodness of fit for each *DSM-IV-TR* and *ICSD-2* insomnia diagnosis listed. These ratings were converted into numeric values reflecting their locations on the 100-pixel VAS and used as the primary data for our multitrait-multimethod analyses.

INTERVIEWER PAIRS AND INTERVIEW PROCESS

Six clinicians at each study site were stratified by sleep medicine experience ($<$10 years vs $\geq$10 years) and by professional degree (MD vs PhD). They were then randomly paired within strata to form 3 pairs who were reasonably similar in their experience and mix of clinical specialties. They were then assigned their respective assessment method: (1) solely the DSISD; (2) a combination of an unstructured clinical interview, sleep history questionnaires, and sleep diaries; or (3) a combination of an unstructured clinical interview, sleep history questionnaires, sleep diaries, and PSG information. All clinicians received training in the use of the computerized diagnostic rating forms, and clinicians using the DSISD also were given training in its administration. Each participant underwent 4 interviews (2 structured and 2 unstructured interviews). Clinicians using the structured sleep interview method conducted separate interviews because the DSISD required independent administration and interpretation. The 2 remaining clinician pairs each conducted a joint interview with each participant. During joint interviews, one clinician interviewed the participant while the other clinician remained silent. When the initial interviewer gained sufficient information to formulate diagnostic impressions, he or she exited the interview room and the second clinician interviewed the participant further if desired. The pair using the unstructured clinical interview and PSG method also reviewed PSG results. A randomization procedure was used so that each clinician served as the initial interviewer for a randomly determined 50% of the participants interviewed.

PROCEDURE

Study candidates first underwent telephone screening with the site's project coordinator. Those passing this screen next met with the project coordinator to provide informed consent and undergo a Mini-Mental State Examination.[34] Those who passed the Mini-Mental State Examination screening were enrolled and completed the following: (1) the Structured Clinical Interview for *DSM-IV* Disorders; (2) the sleep history questionnaire and sleep diary monitoring; and (3) the PSG sessions. Participants then were stratified by sex and age group (aged 18-39, 40-59, and $\geq$60 years) and randomized to 1 of the 6 possible orders of interviews within strata. After completing all interviews, participants chose an in-person or telephone debriefing with the principal investigator (J.D.E. at Duke University Medical Center) or co–principal investigator (E.J.S. or J.K.W. at Rush University Medical Center). During debriefing, the PSG-informed final insomnia diagnosis (or diagnoses) were shared and a treatment referral was made if desired.

ANALYTIC APPROACH

We followed traditional analytic guidelines[28] for multitrait-multimethod research designs. Analyses entailed computing correlations among the clinicians' VAS ratings across methods and diagnoses within each diagnostic system (*DSM-IV-TR* and

ICSD-2) separately. Resulting correlation matrices were used to evaluate reliability and validity for each diagnosis. *Reliability* connotes the degree of agreement between clinicians who use the same assessment method; this is commonly called *interrater reliability*. The correlations between ratings made by the 2 clinicians within each pair for each diagnosis served as reliability indices. With 3 clinician pairs at each of 2 study sites, a total of 6 reliability correlations were derived for each diagnosis. *Convergent validity* connotes how well clinicians using different assessment methods agree in their diagnoses. The CV indices were those correlations reflecting the level of agreement shown for each diagnosis between clinician pairs using differing assessment methods. Concordant with a method described by Campbell and Fiske,[28] the ratings of paired clinicians were first averaged for each diagnosis. We then computed correlations of the resultant averaged ratings of each diagnosis produced by the 3 distinctive clinician pairs. By using this method, we derived 3 CV correlations per diagnosis at each site for a total of 6 such indices for each diagnosis.

Discriminant validity implies that diagnoses are distinctive and can be discriminated. This construct connotes that agreement between clinicians rating the same diagnosis should be notably greater than the agreement observed between clinicians rating distinctive diagnosis. Hence, DV required consideration of correlations of the averaged diagnostic ratings for discrepant diagnoses (within and between clinician pairs). The DV was supported when there was greater correlation between ratings of the same diagnosis derived by different assessment methods than was found for (1) different diagnoses derived by different methods and (2) different diagnoses assessed by the same method. Because the data were not normally distributed, we used Spearman correlation coefficients in all of these analyses.

RESULTS

CLINICIAN TURNOVER

At Duke University Medical Center, the pair using the unstructured clinical interview with access to PSG information experienced staff turnover. At Rush University Medical Center, each clinician pair changed membership; however, 2 of the pairs (the pair using the structured sleep interview method and the pair using the unstructured clinical interview with access to PSG information) retained one of the clinicians for the entire study period. Correlation analyses ignored staffing changes for the following reasons: (1) the information available to the clinicians (eg, PSG or not) was unchanged; (2) one or both members of the pair remained unchanged for 2 clinician pairs at each site; and (3) clinician characteristics remained reasonably stable.

FREQUENCY ANALYSES

We first examined the percentage of cases wherein all 6 interviewers rated each diagnosis as a possible fit (rating >0) as well as the percentage of cases wherein all 6 clinicians viewed each diagnosis as nonapplicable (rating=0). Primary insomnia, insomnia related to another mental disorder, breathing-related sleep disorder, and insomnia due to a general medical condition were the most frequently selected *DSM-IV-TR* diagnoses. The remaining *DSM-IV-TR* diagnoses were rated less frequently but most were assigned ratings higher than 0 by 1 or more clinicians for at least 20% of the cases. Only the diagnosis of no sleep disorder was so infrequent that it was dropped from our reliability and validity analyses.

Many *ICSD-2* categories were rated infrequently and hence were excluded from analyses. Diagnoses retained were psychophysiological insomnia, paradoxical insomnia, idiopathic insomnia, inadequate sleep hygiene, insomnia due to a mental disorder, insomnia due to a medical condition, insomnia due to a drug or substance, obstructive sleep apnea, circadian rhythm sleep disorder–delayed sleep phase type, restless legs syndrome, periodic limb movement disorder, environmental sleep disorder, and other sleep disorder. These were all assigned a rating higher than 0 by 1 or more clinicians in more than 29% of the cases evaluated.

DIAGNOSTIC RELIABILITY

Table 2 and **Table 3** show the reliability indices obtained. The *DSM-IV-TR* categories with the highest interrater reliability were the insomnias related to another mental disorder or due to a medical condition, breathing-related sleep disorder, and circadian rhythm sleep disorder. More modest reliability estimates were noted for alcohol-related sleep disorder and substance-induced sleep disorder. Results for primary insomnia, dyssomnia not otherwise specified, and other sleep disorder were mixed with lower reliability estimates found at Rush University Medical Center.

The *ICSD-2* diagnoses showing the greatest interrater agreement included insomnia due to a mental disorder, insomnia due to a medical condition, periodic limb movement disorder, restless legs syndrome, obstructive sleep apnea, and circadian rhythm sleep disorder–delayed sleep phase type. More modest reliability indices were obtained for the diagnosis of insomnia due to a drug or substance. The Rush University Medical Center site showed lower reliability estimates than the Duke University Medical Center site for psychophysiological insomnia and idiopathic insomnia using the clinical interview method. The Duke University Medical Center site had lower reliability estimates than the Rush University Medical Center site for inadequate sleep hygiene within the clinician pair using the structured sleep interview method and the clinician pair using the unstructured clinical interview with access to PSG information. Interviewers across both sites showed better agreement for the paradoxical insomnia diagnosis when given sleep history questionnaire and diary data to review compared with the DSISD only. The category of other sleep disorder showed much lower reliability at Rush University Medical Center than at Duke University Medical Center.

VALIDITY ANALYSES

For our validity analyses, we retained diagnoses that showed at least modest reliability (mean $r>.20$) and/or had high endorsement rates (rated >0 by ≥ 1 clinician for $\geq 50\%$ of all cases across sites). Accordingly, we eliminated the category of other sleep disorder listed in the *DSM-IV-TR* and *ICSD-2* because of poor reliability at the Rush University Medical Center site and infrequent use overall.

Table 2. Interrater Reliability Indices for *DSM-IV-TR* Insomnia Diagnoses Across Study Sites and Assessment Methods[a]

	Interrater Reliability, % (95% CI)					
	Structured Interview		Clinical Interview		Clinical Interview + PSG	
DSM-IV-TR Insomnia Diagnosis	Duke	Rush	Duke	Rush	Duke	Rush
Primary insomnia	0.44 (0.32 to 0.55)[b]	0.08 (−0.08 to 0.24)	0.43 (0.31 to 0.54)[b]	0.14 (−0.02 to 0.29)	0.28 (0.14 to 0.40)[b]	0.30 (0.14 to 0.44)[b]
Breathing-related sleep disorder	0.76 (0.69 to 0.81)[b]	0.35 (0.20 to 0.48)[b]	0.66 (0.57 to 0.73)[b]	0.61 (0.49 to 0.70)[b]	0.70 (0.62 to 0.76)[b]	0.62 (0.51 to 0.71)[b]
Circadian rhythm sleep disorder	0.44 (0.32 to 0.54)[b]	0.50 (0.37 to 0.61)[b]	0.71 (0.64 to 0.77)[b]	0.46 (0.32 to 0.58)[b]	0.43 (0.31 to 0.54)[b]	0.39 (0.24 to 0.51)[b]
Dyssomnia NOS	0.40 (0.27 to 0.51)[b]	0.33 (0.18 to 0.46)[b]	0.53 (0.42 to 0.62)[b]	−0.004 (−0.17 to 0.16)	0.26 (0.12 to 0.38)[b]	0.08 (−0.08 to 0.24)
Related to a mental disorder	0.59 (0.50 to 0.68)[b]	0.51 (0.38 to 0.62)[b]	0.62 (0.52 to 0.69)[b]	0.60 (0.49 to 0.70)[b]	0.55 (0.45 to 0.64)[b]	0.71 (0.62 to 0.78)[b]
Due to a medical condition	0.43 (0.31 to 0.54)[b]	0.38 (0.23 to 0.51)[b]	0.58 (0.48 to 0.66)[b]	0.47 (0.33 to 0.58)[b]	0.47 (0.35 to 0.57)[b]	0.57 (0.45 to 0.67)[b]
Alcohol-related sleep disorder	0.31 (0.18 to 0.43)[b]	0.39 (0.24 to 0.52)[b]	0.57 (0.47 to 0.66)[b]	0.42 (0.27 to 0.54)[b]	0.21 (0.07 to 0.34)[c]	0.38 (0.23 to 0.51)[b]
Due to a substance	0.24 (0.10 to 0.36)[b]	0.25 (0.10 to 0.40)[c]	0.29 (0.16 to 0.41)[b]	0.40 (0.26 to 0.53)[b]	0.27 (0.14 to 0.39)[b]	0.31 (0.16 to 0.45)[b]
Other sleep disorder	0.22 (0.09 to 0.35)[c]	0.04 (−0.13 to 0.19)	0.23 (0.10 to 0.36)[b]	−0.09 (−0.24 to 0.08)	0.24 (0.10 to 0.36)[b]	0.05 (−0.11 to 0.21)

Abbreviations: CI, confidence interval; NOS, not otherwise specified; PSG, polysomnography.
[a] Duke indicates Duke University Medical Center; Rush, Rush University Medical Center.
[b] $P < .001$.
[c] $P < .01$.

Table 3. Interrater Reliability Indices for *ICSD-2* Insomnia Diagnoses Across Study Sites and Assessment Methods[a]

	Interrater Reliability, % (95% CI)					
	Structured Interview		Clinical Interview		Clinical Interview + PSG	
ICSD-2 Insomnia Diagnosis	Duke	Rush	Duke	Rush	Duke	Rush
Psychophysiological	0.51 (0.40 to 0.60)[b]	0.34 (0.19 to 0.47)[b]	0.52 (0.41 to 0.61)[b]	0.12 (−0.04 to 0.27)	0.27 (0.14 to 0.39)[b]	0.55 (0.42 to 0.65)[b]
Paradoxical	0.15 (0.01 to 0.28)[c]	0.12 (−0.04 to 0.28)	0.51 (0.40 to 0.61)[b]	0.39 (0.24 to 0.52)[b]	0.50 (0.39 to 0.60)[b]	0.41 (0.26 to 0.53)[b]
Idiopathic	0.72 (0.65 to 0.78)[b]	0.57 (0.45 to 0.67)[b]	0.79 (0.73 to 0.84)[b]	0.16 (0.0001 to 0.31)[c]	0.41 (0.29 to 0.52)[b]	0.43 (0.28 to 0.55)[b]
Due to a mental disorder	0.59 (0.49 to 0.67)[b]	0.53 (0.40 to 0.63)[b]	0.63 (0.54 to 0.71)[b]	0.61 (0.50 to 0.70)[b]	0.55 (0.44 to 0.64)[b]	0.67 (0.57 to 0.75)[b]
Due to a medical condition	0.48 (0.37 to 0.58)[b]	0.42 (0.28 to 0.55)[b]	0.63 (0.53 to 0.70)[b]	0.54 (0.42 to 0.65)[b]	0.51 (0.40 to 0.61)[b]	0.57 (0.45 to 0.67)[b]
Due to a drug or substance	0.28 (0.15 to 0.40)[b]	0.30 (0.14 to 0.43)[b]	0.36 (0.24 to 0.48)[b]	0.34 (0.19 to 0.47)[b]	0.31 (0.18 to 0.43)[b]	0.40 (0.26 to 0.53)[b]
Obstructive sleep apnea	0.76 (0.70 to 0.81)[b]	0.49 (0.35 to 0.60)[b]	0.68 (0.59 to 0.74)[b]	0.64 (0.54 to 0.73)[b]	0.66 (0.57 to 0.73)[b]	0.64 (0.53 to 0.72)[b]
Delayed sleep phase syndrome	0.64 (0.55 to 0.72)[b]	0.27 (0.12 to 0.42)[b]	0.75 (0.68 to 0.80)[b]	0.53 (0.40 to 0.63)[b]	0.38 (0.26 to 0.49)[b]	0.40 (0.26 to 0.53)[b]
Restless legs syndrome	0.70 (0.62 to 0.76)[b]	0.41 (0.27 to 0.54)[b]	0.79 (0.73 to 0.83)[b]	0.44 (0.30 to 0.56)[b]	0.69 (0.61 to 0.76)[b]	0.61 (0.50 to 0.70)[b]
Periodic limb movement disorder	0.56 (0.45 to 0.64)[b]	0.56 (0.43 to 0.66)[b]	0.51 (0.40 to 0.61)[b]	0.40 (0.25 to 0.52)[b]	0.54 (0.43 to 0.63)[b]	0.57 (0.45 to 0.67)[b]
Inadequate sleep hygiene	0.13 (−0.01 to 0.26)	0.32 (0.17 to 0.46)[b]	0.47 (0.35 to 0.57)[b]	0.31 (0.16 to 0.45)[b]	0.03 (−0.11 to 0.17)	0.24 (0.08 to 0.38)[d]
Environmental	0.22 (0.09 to 0.35)[d]	0.34 (0.19 to 0.47)[b]	0.34 (0.21 to 0.46)[b]	0.41 (0.27 to 0.53)[b]	0.19 (0.05 to 0.32)[d]	0.46 (0.32 to 0.57)[b]
Other sleep disorder	0.43 (0.31 to 0.53)[b]	0.10 (−0.06 to 0.26)	0.45 (0.33 to 0.55)[b]	0.12 (−0.04 to 0.27)	0.46 (0.34 to 0.56)[b]	−0.01 (−0.17 to 0.15)

Abbreviations: CI, confidence interval; *ICSD-2, International Classification of Sleep Disorders*, second edition; PSG, polysomnography.
[a] Duke indicates Duke University Medical Center; Rush, Rush University Medical Center.
[b] $P < .001$.
[c] $P < .05$.
[d] $P < .01$.

Table 4. Validity Indices for Selected *DSM-IV-TR* Insomnia Subtypes[a]

	SI			CI			CI + PSG		
Insomnia Diagnosis	**SI**	**CI**	**CI + PSG**	**SI**	**CI**	**CI + PSG**	**SI**	**CI**	**CI + PSG**
Duke University Medical Center									
(A) Primary									
CV		.57[b]	.33[b]			.33[b]			
DV	–0.03 (C)	0.01 (G)	0.07 (C)	–0.01 (F)	0.02 (G)	0.08 (F)	–0.01 (E)	0.01 (G)	–0.04 (F)
(B) BRSD									
CV		.77[b]	.46[b]			.43[b]			
DV	0.09 (E)	0.05 (F)	0.18 (H)	0.09 (E)	0.04 (G)	0.16 (E)	0.11 (F)	0.16 (F)	0.08 (F)
(C) CRSD									
CV		.66[b]	.56[b]			.52[b]			
DV	–0.03 (A)	–0.02 (A)	0.10 (H)	–0.01 (A)	–0.01 (D)	0.10 (G)	0.07 (A)	0.03 (A)	0.02 (G)
(D) Mental disorder									
CV		.69[b]	.64[b]			.68[b]			
DV	0.17 (E)	0.23 (E)	0.11 (E)	0.09 (F)	0.18 (E)	0.11 (E)	0.11 (F)	0.14 (E)	0.11 (E)
(E) Medical condition									
CV		.56[b]	.62[b]			.57[b]			
DV	0.17 (D)	0.18 (H)	0.09 (G)	0.23 (D)	0.18 (D)	0.14 (D)	0.12 (B)	0.19 (H)	0.30 (H)
(F) Alcohol induced									
CV		.72[b]	.49[b]			.52[b]			
DV	0.13 (G)	0.09 (D)	0.11 (B)	0.24 (G)	0.10 (E)	0.16 (B)	0.20 (G)	0.09 (E)	0.10 (G)
(G) Substance induced									
CV		.34[b]	.35[b]			.21[c]			
DV	0.13 (F)	0.24 (F)	0.20 (F)	0.07 (D)	0.07 (E)	0.06 (B)	0.09 (E)	0.10 (C)	0.13 (E)
(H) Dyssomnia NOS									
CV		.43[b]	.38[b]			.31[b]			
DV	0.11 (D)	0.07 (D)	0.12 (E)	0.18 (E)	0.16 (E)	0.19 (E)	0.18 (B)	0.11 (E)	0.30 (E)
Rush University Medical Center									
(A) Primary		NR	NR			NR			
(B) BRSD									
CV		.58[b]	.51[b]			.34[b]			
DV	0.03 (F)	0.02 (F)	0.09 (E)	0.12 (E)	0.08 (F)	0.15 (E)	0.09 (E)	0.05 (E)	0.11 (E)
(C) CRSD									
CV		.45[b]	.50[b]			.39[b]			
DV	0.15 (G)	0.14 (G)	0.05 (B)	0.10 (G)	0.16 (F)	0.07 (G)	0.09 (G)	0.14 (D)	0.05 (G)
(D) Mental disorder									
CV		.61[b]	.73[b]			.64[b]			
DV	0.22 (G)	0.18 (G)	0.23 (G)	0.23 (F)	0.12 (F)	0.14 (C)	0.20 (F)	0.20 (G)	0.23 (G)
(E) Medical condition									
CV		.50[b]	.58[b]			.45[b]			
DV	0.19 (G)	0.12 (B)	0.13 (D)	0.07 (G)	0.05 (F)	0.05 (B)	0.11 (G)	0.15 (B)	0.18 (G)
(F) Alcohol induced									
CV		.36[b]	.30[b]			.42[b]			
DV	0.20 (D)	0.38 (G)	0.20 (D)	0.11 (C)	0.26 (G)	0.15 (D)	0.19 (D)	0.25 (G)	0.19 (G)
(G) Substance induced									
CV		.31[b]	.49[b]			.39[b]			
DV	0.22 (D)	0.10 (C)	0.17 (D)	0.39 (F)	0.26 (F)	0.25 (F)	0.23 (D)	0.12 (F)	0.23 (D)
(H) Dyssomnia NOS									
CV		NR	NR			NR			

Abbreviations: BRSD, breathing-related sleep disorder; CI, unstructured clinical interview method; CI + PSG, unstructured clinical interview with access to polysomnography information; CRSD, circadian rhythm sleep disorder; CV, convergent validity; DV, discriminant validity; NOS, not otherwise specified; NR, not reported owing to low reliability at study site indicated; SI, structured sleep interview method.

[a] The CVs for the SI vs CI, SI vs CI + PSG, and CI vs CI + PSG methods are listed only once for each diagnosis to eliminate redundancy. The DVs shown are the highest correlations derived from the monomethod-heterotrait comparisons (ie, comparisons for differing diagnoses within methods of assessment) and the heteromethod-heterotrait comparisons (ie, comparisons for differing diagnoses across differing methods of assessment). For the DVs, adjacent letters in parentheses connote the diagnosis most strongly correlated with the diagnosis listed in the first column. For example, "–0.03 (C)" shown in the third column of the first row of the Duke University Medical Center data indicates the correlation between the target diagnosis (primary insomnia) and CRSD derived within the structured interview method.

[b] $P < .001$.

[c] $P < .01$.

Table 4 and **Table 5** show CV and DV indices derived for the *DSM-IV-TR* and *ISCD-2* insomnia diagnoses examined. A diagnosis is considered valid when the CV values are statistically significant (with higher values connoting greater CV) and the DV values are consistently lower than the CV values and, preferably, nonsignificant. The CV and DV values are not reported when poor reliability was found for that diagnosis within a study site.

Table 5. Validity Indices for Selected *ICSD-2* Insomnia Subtypes[a]

	SI			CI			CI + PSG		
Insomnia Diagnosis	**SI**	**CI**	**CI + PSG**	**SI**	**CI**	**CI + PSG**	**SI**	**CI**	**CI + PSG**
				Duke University Medical Center					
(A) Psychophysiological									
CV		.50[b]	.42[b]			.27[b]			
DV	0.03 (L)	0.06 (H)	0.11 (K)	0.10 (D)	0.11 (K)	0.18 (K)	0.08 (F)	0.04 (K)	0.06 (C)
(B) Paradoxical									
CV		.24[b]	.14			.22[c]			
DV	0.22 (C)	0.21 (C)	0.28 (C)	0.12 (C)	0.15 (C)	0.11 (C)	0.24 (K)	0.12 (K)	0.15 (C)
(C) Idiopathic									
CV		.72[b]	.58[b]			.68[b]			
DV	0.24 (H)	0.26 (H)	0.19 (H)	0.31 (H)	0.30 (H)	0.21 (H)	0.28 (B)	0.11 (B)	0.15 (B)
(D) Mental disorder									
CV		.69[b]	.66[b]			.66[b]			
DV	0.18 (F)	0.19 (F)	0.12 (F)	0.09 (I)	0.12 (F)	0.08 (F)	0.08 (I)	0.13 (F)	0.10 (I)
(E) Inadequate sleep hygiene									
CV		NR	NR			NR			
(F) Medical disorder									
CV		.57[b]	.65[b]			.57[b]			
DV	0.18 (I)	0.20 (I)	0.17 (J)	0.19 (D)	0.29 (K)	0.13 (D)	0.25 (I)	0.30 (I)	0.23 (I)
(G) OSA									
CV		.74[b]	.45[b]			.38[b]			
DV	0.21 (L)	0.14 (L)	0.13 (L)	0.13 (F)	0.13 (L)	0.18 (F)	0.10 (L)	0.09 (I)	0.12 (K)
(H) DSPS									
CV		.72[b]	.59[b]			.53[b]			
DV	0.24 (C)	0.31 (C)	0.08 (I)	0.26 (C)	0.30 (C)	0.06 (J)	0.19 (C)	0.21 (C)	0.13 (C)
(I) RLS									
CV		.63[b]	.72[b]			.68[b]			
DV	0.43 (L)	0.41 (L)	0.26 (L)	0.34 (L)	0.50 (L)	0.30 (F)	0.42 (L)	0.41 (L)	0.23 (F)
(J) Due to drug or substance									
CV		.47[b]	.37[b]			.37[b]			
DV	0.14 (F)	0.18 (K)	0.15 (F)	0.05 (D)	0.12 (K)	0.13 (L)	0.17 (F)	0.10 (F)	0.17 (F)
(K) Environmental									
CV		.30[b]	.25[b]			.31[b]			
DV	0.14 (L)	0.14 (I)	0.24 (B)	0.18 (J)	0.29 (F)	0.12 (F)	0.11 (A)	0.18 (A)	0.12 (G)
(L) Periodic limb movements									
CV		.55[b]	.29[b]			.23[c]			
DV	0.43 (I)	0.34 (I)	0.42 (I)	0.41 (I)	0.50 (I)	0.41 (I)	0.26 (I)	0.13 (J)	0.14 (I)

(continued)

The best-supported *DSM-IV-TR* diagnosis was insomnia associated with another mental disorder (Table 4). Insomnia due to a medical condition, breathing-related sleep disorder, and circadian rhythm sleep disorder also showed reasonable validity. Alcohol-related sleep disorder showed more modest validity indices, whereas other substance-induced insomnia received less support. Primary insomnia and dyssomnia not otherwise specified were least supported. Their CV indices were in the low to medium range across methods at Duke University Medical Center; validity correlations for primary insomnia and dyssomnia not otherwise specified were not calculated at Rush University Medical Center owing to low reliability indices noted there.

The best-supported *ICSD-2* diagnoses were insomnia due to a mental disorder and insomnia due to a medical condition (Table 5); their CV indices mainly fell in the large range and the DV indices were generally in the insignificant range. Obstructive sleep apnea was also reasonably supported, but its pattern of correlations suggested that clinicians who had access to PSG information differed from those who did not. Restless legs syndrome, circadian rhythm sleep disorder–delayed sleep phase type, and idiopathic insomnia received more modest support: the CV indices for these categories fell in the medium to large range and were generally larger than their related DV indices. Insomnia due to a drug or substance and environmental sleep disorder received less support, with the CV indices falling in the small to medium range. At Duke University Medical Center, the validity indices for psychophysiological insomnia ranged from small to large, but poor reliability for this diagnosis at Rush University Medical Center obviated its validity testing there. A comparison of CV and DV indices acquired for inadequate sleep hygiene suggested reasonable validity for this diagnosis at the Rush University Medical Center site, but its poor reliability at Duke University Medical Center prevented assessing its validity there. The validity indices were variable for periodic limb movement disorder across sites and methods, whereas paradoxical insomnia received little support overall.

Table 5. Validity Indices for Selected *ICSD-2* Insomnia Subtypes[a] (continued)

	SI			CI			CI + PSG		
Insomnia Diagnosis	**SI**	**CI**	**CI + PSG**	**SI**	**CI**	**CI + PSG**	**SI**	**CI**	**CI + PSG**
				Rush University Medical Center					
(A) Psychophysiological									
CV		NR	NR			NR			
(B) Paradoxical									
CV		.07	.18[d]			.19[d]			
DV	0.16 (C)	0.31 (C)	0.18 (C)	0.12 (F)	0.02 (F)	0.13 (C)	0.05 (E)	0.08 (H)	0.16 (C)
(C) Idiopathic									
CV		.33[b]	.57[b]			.32[b]			
DV	0.16 (B)	0.21 (D)	0.20 (H)	0.31 (B)	0.25 (H)	0.20 (L)	0.18 (B)	0.20 (I)	0.18 (I)
(D) Mental disorder									
CV		.66[b]	.69[b]			.64[b]			
DV	0.25 (J)	0.18 (I)	0.18 (J)	0.21 (C)	0.14 (L)	0.16 (H)	0.20 (J)	0.17 (J)	0.28 (J)
(E) Inadequate sleep hygiene									
CV		.34[b]	.28[b]			.37[b]			
DV	0.31 (G)	0.13 (K)	0.13 (G)	0.19 (J)	0.17 (J)	0.16 (J)	0.20 (K)	0.26 (J)	0.16 (K)
(F) Medical disorder									
CV		.53[b]	.57[b]			.46[b]			
DV	0.29 (I)	0.20 (I)	0.17 (J)	0.20 (I)	0.32 (L)	0.25 (I)	0.09 (K)	0.17 (G)	0.17 (J)
(G) OSA									
CV		.53[b]	.50[b]			.30[b]			
DV	0.31 (E)	0.13 (E)	0.23 (J)	0.16 (I)	0.32 (L)	0.40 (L)	0.13 (E)	0.12 (B)	0.10 (J)
(H) DSPS									
CV		.37[b]	.50[b]			.38[b]			
DV	0.24 (J)	0.19 (C)	0.16 (C)	0.13 (J)	0.25 (C)	0.17 (E)	0.20 (C)	0.16 (D)	0.17 (C)
(I) RLS									
CV		.36[b]	.34[b]			.50[b]			
DV	0.29 (F)	0.20 (F)	0.13 (D)	0.20 (F)	0.54 (L)	0.36 (L)	0.15 (F)	0.25 (F)	0.24 (L)
(J) Due to drug or substance									
CV		.24[c]	.47[b]			.22[c]			
DV	0.25 (D)	0.19 (E)	0.20 (H)	0.13 (E)	0.17 (E)	0.26 (E)	0.23 (G)	0.16 (E)	0.28 (D)
(K) Environmental									
CV		.31[b]	.39[b]			.39[b]			
DV	0.15 (H)	0.23 (L)	0.21 (L)	0.13 (E)	0.12 (E)	0.19 (E)	0.04 (B)	0.09 (L)	0.16 (E)
(L) Periodic limb movements									
CV		.27[b]	.10			.48[b]			
DV	0.28 (I)	0.13 (G)	0.09 (F)	0.23 (K)	0.54 (I)	0.21 (I)	0.21 (K)	0.40 (G)	0.24 (I)

Abbreviations: CI, unstructured clinical interview method; CI + PSG, unstructured clinical interview with access to polysomnography information; CV, convergent validity; DSPS, delayed sleep phase syndrome; DV, discriminant validity; *ICSD-2*, *International Classification of Sleep Disorders*, second edition; NR, not reported owing to low reliability at study site indicated; OSA, obstructive sleep apnea; RLS, restless legs syndrome; SI, structured sleep interview method.

[a] The CVs for the SI vs CI, SI vs CI + PSG, and CI vs CI + PSG methods are listed only once for each diagnosis to eliminate redundancy. The DVs shown are the highest correlations derived from the monomethod-heterotrait comparisons (ie, comparisons for differing diagnoses within methods of assessment) and the heteromethod-heterotrait comparisons (ie, comparisons for differing diagnoses across differing methods of assessment). For the DVs, adjacent letters in parentheses connote the diagnosis most strongly correlated with the diagnosis listed in the first column. For example, "0.03 (L)" shown in the third column of the first row of the Duke University Medical Center data indicates the correlation between the target diagnosis (psychophysiological insomnia) and periodic limb movements derived within the structured interview method.

[b] $P < .001$.

[c] $P < .01$.

[d] $P < .05$.

DATA SYNTHESIS

Because the study design and data obtained may be novel to many readers, we conducted an additional classification analysis to summarize results and place them in a practical context. Since to our knowledge there are no standardized methods for classifying outcomes of multitrait-multimethod studies, we offer the rationally derived classification rules in **Table 6**. These rules consider the size and significance of the reliability and validity correlations obtained to gauge the acceptability of each diagnosis. The correlations themselves were appraised using Cohen's guidelines,[35] wherein correlation coefficients in the order of 0.10 are regarded as small, those of 0.30 are medium, and those of 0.50 or higher are large. These cutoffs could be considered arbitrary when applied to individual reliability or validity indices, but our classification approach arguably provides a practical manner for synthesizing our results.

Table 7 shows how the diagnoses rate when applying these classification rules. Within *DSM-IV-TR*, insomnia related to another mental disorder is rated as highly acceptable; breathing-related sleep disorder, insomnia due to a medical condition, and circadian rhythm sleep disorder are acceptable; alcohol-related sleep disorder is marginally acceptable; and the remaining diagnoses are unacceptable. Within *ICSD-2*, insomnia due to a mental disorder is highly acceptable; obstructive sleep apnea and insomnia due to a

Table 6. Classification System[a]

Classification	Reliability	CV	DV
Highly acceptable	All *r* values are in large range (>0.50)	All CV *r* values are in large range (>0.50)	All DV *r* values are in small range (<0.30) or insignificant range
Acceptable	All *r* values are in medium range (>0.30) or high range (>0.50)	All CV *r* values are in medium range (>0.30) or high range (>0.50)	>80% of DV *r* values are in small range (≤0.30) or insignificant range
Marginally acceptable	All *r* values are significant and ≤1 *r* value falls in small range (<0.30); remainder fall in medium or large range	All CV *r* values are significant and ≤1 *r* value falls in small range (<.30); remainder fall in medium or large range	DV *r* values are consistently lower than CV *r* values with ≤1 DV greater than or equal to site-specific CV values obtained
Unacceptable	Does not meet any above criteria	Does not meet any above criteria	Does not meet any above criteria

Abbreviations: CV, convergent validity; DV, discriminant validity.
[a] The classification system for correlation coefficients is by Cohen.[35]

Table 7. Classification Results

Classification	*DSM-IV-TR* Diagnoses	*ICSD-2* Diagnoses
Highly acceptable	Related to a mental disorder	Due to a mental disorder
Acceptable	Breathing-related sleep disorder; due to a medical condition; circadian rhythm sleep disorder	Obstructive sleep apnea; due to a medical condition
Marginally acceptable	Alcohol-related sleep disorder	Restless legs syndrome; delayed sleep phase type; idiopathic
Unacceptable	Primary insomnia; dyssomnia NOS; substance-induced sleep disorder; other sleep disorder	Psychophysiological; paradoxical; due to a drug or substance; periodic limb movement disorder; inadequate sleep hygiene; environmental sleep disorder; other sleep disorder

Abbreviations: *ICSD-2*, *International Classification of Sleep Disorders*, second edition; NOS, not otherwise specified.

medical condition are acceptable; restless legs syndrome, circadian rhythm sleep disorder–delayed sleep phase type, and idiopathic insomnia are marginally acceptable; and the remaining diagnoses are unacceptable.

COMMENT

The *DSM-IV-TR* and *ICSD-2* sleep disorders nosologies are used widely, but few studies have tested their reliability and validity. Results of this trial show that each system includes diagnoses with acceptable reliability and validity. Within *DSM-IV-TR*, insomnia related to another mental disorder, insomnia due to a general medical condition, breathing-related sleep disorder, and circadian rhythm sleep disorder were best supported. Alcohol-related sleep disorder was more marginally supported, whereas other substance-induced sleep disorder was not well supported. Least supported were dyssomnia not otherwise specified, other sleep disorder, and, surprisingly, primary insomnia. Although primary insomnia was frequently rated, our data call into question its reliability and validity.

Within *ICSD-2*, insomnia due to a mental disorder, insomnia due to a medical condition, and obstructive sleep apnea were well supported. Circadian rhythm sleep disorder–delayed sleep phase type and idiopathic insomnia also received reasonable support and fell in the marginally acceptable classification. Restless legs syndrome was also classified as marginally acceptable rather than acceptable largely owing to its overlap or correlation with periodic limb movement disorder. However, it is well recognized that periodic limb movements are highly prevalent among patients with restless legs syndrome.[36] Given this consideration, perhaps the restless legs syndrome diagnosis should be classified as acceptable. In contrast, insomnia due to a drug or substance and environmental sleep disorder appeared much more marginal. Support for the remaining *ICSD-2* diagnoses was more variable and generally poor.

Table 7 suggests that an optimal insomnia nosology derives from a melding of *DSM-IV-TR* and *ICSD-2* diagnoses. The categories of insomnia related or due to a mental disorder, insomnia due to a medical condition, and breathing-related sleep disorder or obstructive sleep apnea seemingly merit strong consideration for inclusion. These categories occur within both nosologies but have slightly different labels in each. Restless legs syndrome also seems to be a viable diagnosis, as does *DSM-VI-TR* circadian rhythm sleep disorder, which seems favored over the more specific *ICSD-2* delayed sleep phase type. Finally, *DSM-IV-TR* alcohol-related sleep disorder along with *ICSD-2* idiopathic insomnia merit consideration as well.

The *DSM-IV-TR* diagnosis of primary insomnia and most related *ICSD-2* subtypes were rated frequently by study clinicians but garnered minimal support. The correlations obtained suggest that the addition of PSG data to interview findings seemingly complicates diagnostic ascertainment and reduces reliability and validity for most of these categories. Because cognitive and behavioral mechanisms are thought to be important perpetuating mechanisms in primary insomnia and its related *ICSD-2* subtypes, assignment of these diagnoses is largely dependent on ascertaining the presence of such mechanisms as the likely cause of the insomnia disorder ob-

served. However, patients presumed to have primary insomnia often have some symptoms of depression and/or anxiety, whereas those presumed to have comorbid forms of insomnia often manifest the cognitive-behavioral aberrations thought to perpetuate primary insomnia.[37] Hence, primary vs comorbid insomnia distinctions are often subtle and perhaps even arbitrary. Perhaps refinement of the definitional criteria for primary insomnia and its related *ICSD-2* subtypes could improve their reliability and validity, but efforts to do so will be complicated by their degree of overlap with the comorbid insomnia subtypes. Ultimately, it may be necessary to adopt a different paradigm for insomnia diagnosis.

Perhaps the primary vs comorbid insomnia distinction could be abandoned in favor of a more inclusive diagnostic term such as *insomnia disorder*. This is the approach being taken in the upcoming *DSM-5*. Patients who meet the insomnia criteria outlined in the Research Diagnostic Criteria[26] for insomnia disorder are to be assigned this diagnosis regardless of any coincident sleep-disruptive comorbidities. However, diagnosticians are encouraged to also diagnose coexisting psychiatric and medical comorbidities. Such an approach should simplify clinicians' diagnostic task with patients who have insomnia. Of course, use of the global insomnia disorder diagnosis could also encourage a generic one-size-fits-all approach to insomnia treatment. Whether this concern is warranted will only be determined once the *DSM-5* is placed into use.

We could also consider abandoning the sorting of patients into diagnostic "bins" but instead using dimensional measures of insomnia symptoms for patient characterization. This approach is advocated by the developers of Profile Analysis via Multidimensional Scaling.[38,39] This method applies a multidimensional scaling analysis to patients' scores on syndrome-relevant questionnaires to identify core symptom profiles. Profile Analysis via Multidimensional Scaling then locates each patient's symptom pattern by assigning patient-specific weights for each profile. These weights connote the degree of match with each core profile and designate the patient's exact diagnostic location in relation to these profiles. Patients are not forced into single, often poorly fitting diagnostic categories but rather are characterized by their overall symptom arrays. This method may more accurately characterize each patient and support efforts toward the development of individualized therapies. The *DSM-5* advocates use of dimensional measures, so perhaps future nosologies will incorporate this method.

In deciding on future changes to our insomnia diagnostic systems, it may be useful to consider this study's findings in conjunction with other projects such as the field trials for the *DSM-5* sleep disorders nosology. Our data provide information about interrater reliability and the validity and credibility of the insomnia diagnoses examined in the eyes of our study clinicians. However, the *DSM-5* trials will include many more study sites than used in this project and will produce the test-retest reliability data that were not provided with our method. Such additional findings in conjunction with our results should guide future improvements to our insomnia nosologies.

It is important to consider the potential effects of our study method on the findings obtained. Given the statistical demands of the multitrait-multimethod research design, it was necessary to have clinicians rate the goodness of fit of each insomnia diagnosis on VASs for each patient they evaluated. Such ratings may reveal clinicians' subjective decision-making processes when assigning diagnoses, but they do not replicate diagnostic outcomes in clinical situations wherein clinicians assign diagnoses in an all-or-none fashion. Cross-validation of our results using methods that replicate real-world diagnostic practices may therefore be warranted.

Admittedly, this study had several limitations. Our sample included mainly research volunteers. Whether similar results would have been obtained with clinical patients or those with insomnia who are selected from the community remains unknown. Furthermore, data were obtained from only 2 study sites. Replication of this study across additional sites would have been desirable, albeit quite costly. Because many of the subtypes considered were seldom or never rated by study clinicians, the study may have benefited from a much larger and diverse sample. Additionally, the greater turnover in our clinician raters at Rush University Medical Center may have contributed to greater variability in results noted there. Finally, our findings suggested notable differences in the reliability and validity indices obtained across sites for selected diagnoses. Whether such differences should be attributed to demographic differences in their study samples, general site-specific biases, and/or idiosyncratic diagnostic propensities of clinicians cannot be determined given our study design. Nonetheless, few previous studies have assessed reliability and validity of insomnia diagnoses routinely used in practice. Thus, results presented herein fill a void and provide guidance for revisions of our insomnia classification schemes.

Submitted for Publication: October 29, 2010; final revision received March 15, 2011; accepted March 31, 2011.
Published Online: June 6, 2011. doi:10.1001/archgenpsychiatry.2011.64
Author Affiliations: Veterans Affairs Medical Center (Drs Edinger, Olsen, and Means and Ms Stechuchak) and Duke University Medical Center (Drs Edinger, Olsen, Chiang, Lineberger, Means, Radtke, and Krystal), Durham, North Carolina; Sleep Disorders Service and Research Center, Department of Behavioral Sciences, Rush University Medical Center, Chicago, Illinois (Drs Wyatt and Crisostomo); ACORN Research, LLC, Memphis, Tennessee (Dr Stepanski); Ryerson University, Toronto, Ontario, Canada (Dr Carney); and Veterans Affairs Medical Center, Miami, Florida (Dr Wohlgemuth).
Correspondence: Jack D. Edinger, PhD, Psychology Service (116B), Veterans Affairs Medical Center, 508 Fulton St, Durham, NC 27705 (jack.edinger@duke.edu).
Financial Disclosure: Dr Edinger has been a consultant for Kingsdown and Philips/Respironics and has received research support from Philips/Respironics. Dr Wyatt has received research support from Philips/Respironics. Dr Krystal has received grants or research support from the National Institutes of Health, sanofi-aventis, Cephalon, GlaxoSmithKline, Merck, Neurocrine, Pfizer, Sepracor,

Somaxon, Takeda, Transcept, Philips/Respironics, Neurogen, Evotec, Astellas, and Neuronetics and has been a consultant for Abbott, Actelion, Arena, Astellas, Axiom, AstraZeneca, Bristol-Myers Squibb, Cephalon, Eli Lilly, GlaxoSmithKline, Jazz, Johnson & Johnson, King, Merck, Neurocrine, Neurogen, Neuronetics, Novartis, Organon, Ortho-McNeil-Janssen, Pfizer, Respironics, Roche, sanofi-aventis, Sepracor, Somaxon, Takeda, Transcept, and Kingsdown.

Funding/Support: This research was supported by grant R01 MH67057 from the National Institute of Mental Health.

Role of the Sponsor: The funding agency had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

Disclaimer: The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

Previous Presentation: This paper was presented in part at the 20th Congress of the European Sleep Research Society; September 17, 2010; Lisbon, Portugal.

Additional Contributions: Kevan VanLandingham, MD, PhD, Laurie Keefer, PhD, Andrea Canada, PhD, Aimee Danielson, PhD, Babak Mokhlesi, MD, Margaret Park, MD, Mike Summers, MD, and Tony Proske, MD, served as clinical interviewers and Marci Loiselle, Angela Kirby, Pamela Smith, Faye Knauss, Kathy Schelble, Laura Benson, and Lindsey Gluszek served as study coordinators.

REFERENCES

1. Association of Sleep Disorders Centers; Association for the Psychophysiological Study of Sleep. Diagnostic classification of sleep and arousal disorders: 1979 first edition. *Sleep*. 1979;2(1):1-154.
2. World Health Organization. *International Classification of Diseases, Ninth Revision, Clinical Modification*. Geneva, Switzerland: World Health Organization; 1992.
3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed. Washington, DC: American Psychiatric Association; 1994.
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed, text revision. Washington, DC: American Psychiatric Association; 2000.
5. American Sleep Disorders Association. *International Classification of Sleep Disorders*. Rochester, MN: American Sleep Disorders Association; 1990.
6. American Sleep Disorders Association. *International Classification of Sleep Disorders, Revised: Diagnostic and Coding Manual*. Rochester, MN: American Sleep Disorders Association; 1997.
7. Buysse DJ, Young T, Edinger JD, Carroll J, Kotagal S. Clinicians' use of the International Classification of Sleep Disorders: results of a national survey. *Sleep*. 2003;26(1):48-51.
8. American Academy of Sleep Medicine. *International Classification of Sleep Disorders*. 2nd ed. Darien, IL: American Academy of Sleep Medicine; 2005.
9. Soldatos CR, Kales JD, Tan TL, Kales A. Classification of sleep disorders. *Psychiatr Ann*. 1987;17:454-457.
10. Buysse DJ, Reynolds CF III, Monk TH, Hoch CC, Yeager AL, Kupfer DJ. Quantification of subjective sleep quality in healthy elderly men and women using the Pittsburgh Sleep Quality Index (PSQI). *Sleep*. 1991;14(4):331-338.
11. Vgontzas AN, Kales A, Bixler EO, Vela-Bueno A. In discussion of: Nofzinger EA, Buysse DJ, Reynolds CF 3rd, Kupfer DJ. Sleep disorders related to another mental disorder (nonsubstance/primary): a *DSM-IV* literature review. *J Clin Psychiatry*. 1993;54(7):244-259.
12. Hauri PJ, Olmstead E. Childhood-onset insomnia. *Sleep*. 1980;3(1):59-65.
13. Nofzinger EA, Buysse DJ, Reynolds CF III, Kupfer DJ. Sleep disorders related to another mental disorder (nonsubstance/primary): a *DSM-IV* literature review. *J Clin Psychiatry*. 1993;54(7):244-259.
14. Hauri PJ, Wisbey J. Wrist actigraphy in insomnia. *Sleep*. 1992;15(4):293-301.
15. Edinger JD, Krystal AD. Subtyping primary insomnia: is sleep state misperception a distinct clinical entity? *Sleep Med Rev*. 2003;7(3):203-214.
16. Krystal AD, Edinger JD, Wohlgemuth WK, Marsh GR. NREM sleep EEG frequency spectral correlates of sleep complaints in primary insomnia subtypes. *Sleep*. 2002;25(6):630-640.
17. Downey R III, Bonnet MH. Training subjective insomniacs to accurately perceive sleep onset. *Sleep*. 1992;15(1):58-63.
18. McCall WV, Edinger JD. Subjective total insomnia: an example of sleep state misperception. *Sleep*. 1991;15(1):71-73.
19. Ohayon MM, Roberts RE. Comparability of sleep disorders diagnoses using *DSM-IV* and *ICSD* classifications with adolescents. *Sleep*. 2001;24(8):920-925.
20. Schramm E, Hohagen F, Grasshoff U, Riemann D, Hajak G, Weess HG, Berger M. Test-retest reliability and validity of the Structured Interview for Sleep Disorders According to *DSM-III-R*. *Am J Psychiatry*. 1993;150(6):867-872.
21. Edinger JD, Fins AI, Goeke JM, McMillan DK, Gersh TL, Krystal AD, McCall WV. The empirical identification of insomnia subtypes: a cluster analytic approach. *Sleep*. 1996;19(5):398-411.
22. Morin CM, Stone J, McDonald K, Jones S. Psychological management of insomnia: a clinical replication series with 100 patients. *Behav Ther*. 1994;25:291-309.
23. Buysse DJ, Reynolds CF III, Hauri PJ, Roth T, Stepanski EJ, Thorpy MJ, Bixler EO, Kales A, Manfredi RL, Vgontzas AN, Stapf DM, Houck PR, Kupfer DJ. Diagnostic concordance for *DSM-IV* sleep disorders: a report from the APA/NIMH *DSM-IV* field trial. *Am J Psychiatry*. 1994;151(9):1351-1360.
24. Ohayon MM, Guilleminault C, Zulley J, Palombini L, Raab H. Validation of the sleep-EVAL system against clinical assessments of sleep disorders and polysomnographic data. *Sleep*. 1999;22(7):925-930.
25. Hauri PJ. A cluster analysis of insomnia. *Sleep*. 1983;6(4):326-338.
26. Edinger JD, Bonnet MH, Bootzin RR, Doghramji K, Dorsey CM, Espie CA, Jamieson AO, McCall WV, Morin CM, Stepanski EJ; American Academy of Sleep Medicine Work Group. Derivation of research diagnostic criteria for insomnia: report of an American Academy of Sleep Medicine Work Group. *Sleep*. 2004;27(8):1567-1596.
27. Buysse DJ, Reynolds CF III, Kupfer DJ, Thorpy MJ, Bixler E, Kales A, Manfredi R, Vgontzas A, Stepanski E, Roth T, Hauri P, Stapf D. Effects of diagnosis on treatment recommendations in chronic insomnia: a report from the APA/NIMH *DSM-IV* field trial. *Sleep*. 1997;20(7):542-552.
28. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull*. 1959;56(2):81-105.
29. Rechtshaffen A, Kales A. *A Manual of Standardized Terminology, Techniques, and Scoring Systems of Sleep Stages of Human Subjects*. Bethesda, MD: US National Institute of Neurological Diseases and Blindness, Neurological Information Network; 1968.
30. Coleman R. Periodic movements in sleep (nocturnal myoclonus) and restless legs syndrome. In: Guilleminault C, ed. *Sleeping and Waking Disorders: Indications and Techniques*. Menlo Park, CA: Addison-Wesley; 1982:265-295.
31. American Sleep Disorders Association. EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. *Sleep*. 1992;15(2):173-184.
32. Carney CE, Edinger JD, Olsen MK, Stechuchak KM, Krystal AD, Wyatt JK. Interrater reliability for insomnia diagnoses derived from the Duke Structured Interview for Sleep Disorders. *Sleep*. 2008;31(suppl):A250.
33. Edinger JD, Wyatt JK, Olsen MK, Stechuchak KM, Carney CE, Chiang A, Krystal AD, Lineberger MD, Means MK, Radtke RA, Knauss F. Reliability and validity of insomnia diagnoses derived from the Duke Structured Interview for Sleep Disorders. *Sleep*. 2009;32(suppl):A265.
34. Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198.
35. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.
36. Becker P. Restless legs syndrome. In: Lee-Chiong T, ed. *Sleep: A Comprehensive Handbook*. Hoboken, NJ: John Wiley & Sons; 2006:473-481.
37. Nowell PD, Buysse DJ, Reynolds CF III, Hauri PJ, Roth T, Stepanski EJ, Thorpy MJ, Bixler E, Kales A, Manfredi RL, Vgontzas AN, Stapf DM, Houck PR, Kupfer DJ. Clinical factors contributing to the differential diagnosis of primary insomnia and insomnia related to mental disorders. *Am J Psychiatry*. 1997;154(10):1412-1416.
38. Davison ML, Gasser M, Ding S. Identifying major profile patterns in a population: an exploratory study of WAIS and GATB patterns. *Psychol Assess*. 1996;8(1):26-31.
39. Kim SK, Frisby CL, Davison ML. Estimating cognitive profiles using Profile Analysis via Multidimensional Scaling (PAMS). *Multivariate Behav Res*. 2004;39(4):595-624.