Sensitivity and Specificity of Polysomnographic Criteria for Defining Insomnia

Jack D. Edinger, Ph.D.; Christi S. Ulmer, Ph.D.; Melanie K. Means, Ph.D.

VA and Duke University Medical Centers, Durham, NC

Study Objectives: In recent years, polysomnography-based eligibility criteria have been increasingly used to identify candidates for insomnia research, and this has been particularly true of studies evaluating pharmacologic therapy for primary insomnia. However, the sensitivity and specificity of PSG for identifying individuals with insomnia is unknown, and there is no consensus on the criteria sets which should be used for participant selection. In the current study, an archival data set was used to test the sensitivity and specificity of PSG measures for identifying individuals with primary insomnia in both home and lab settings. We then evaluated the sensitivity and specificity of the eligibility criteria employed in a number of recent insomnia trials for identifying primary insomnia sufferers in our sample.
Design: Archival data analysis.
Settings: Study participants' homes and a clinical sleep laboratory.

Participants: Adults: 76 with primary insomnia and 78 non-complaining normal sleepers.
Measurements and Results: ROC and cross-tabs analyses were used to evaluate the sensitivity and specificity of PSG-derived total sleep time, latency to persistent sleep, wake after sleep onset, and sleep efficiency for discriminating adults with primary insomnia from normal sleepers. None of the individual criteria accurately discriminated PI from normal sleepers, and none of the criteria sets used in recent trials demonstrated acceptable sensitivity and specificity for identifying primary insomnia.
Conclusions: The use of quantitative PSG-based selection criteria in insomnia research may exclude many who meet current diagnostic criteria for an insomnia disorder.
Keywords: Insomnia, PSG, research, sensitivity, specificity
Citation: Edinger JD; Ulmer CS; Means MK. Sensitivity and specificity of polysomnographic criteria for defining insomnia. *J Clin Sleep Med* 2013;9(5):481-491.

Insomnia is a highly prevalent and often debilitating condition that reduces quality of life, increases risks for various mental and medical disorders and enhances healthcare costs for millions of individuals world-wide. Over the past several decades a burgeoning body of insomnia-focused research has advanced our understanding and ability to manage this condition. However, this body of research has not been without its limitations. As noted by Buysse et al.,[1] the insomnia research literature has been plagued by lack of standardization, particularly in regard to the methods and criteria used for assessing and characterizing insomnia symptoms, and perhaps more importantly, the more global insomnia disorder. This lack of standardization, in turn, has resulted in considerable variability in the selection of insomnia samples across studies. Consequently, comparisons of findings across studies are often difficult or impossible to conduct. This state of affairs has slowed our progress toward developing optimal insomnia management strategies and reducing the overall prevalence and public impact of this troublesome sleep disorder.

Fortunately, in recent years, there have been a number of efforts in the sleep research and professional community to remedy this situation. Included among these are the development and publication of research diagnostic criteria for defining insomnia,[2] as well as the publication of consensus recommendations for standardizing assessment methods and measures in all insomnia research studies.[1] Additionally, some investigators have proposed developing quantitative insomnia criteria based on indices of nocturnal sleep disturbance. Studies devoted to this objective[3,4] have

BRIEF SUMMARY

Current Knowledge/Study Rationale: In recent years, polysomnography-based eligibility criteria have been increasingly used to identify candidates for insomnia research. The current study was conducted to evaluate the accuracy of this approach for identifying individuals with insomnia.
Study Impact: We found that none of the PSG-based criteria accurately discriminated those with primary insomnia from normal sleepers, and none of the criteria sets used in recent trials demonstrated acceptable sensitivity and specificity for identifying primary insomnia. Thus, our findings suggest that the use of quantitative PSG-based selection criteria in insomnia research may exclude many who meet current diagnostic criteria for an insomnia disorder.

shown that a cutoff value > 30 minutes for sleep onset latency or wake time after onset derived from respondents' sleep diaries has good sensitivity and specificity for discriminating insomnia sufferers from normal sleepers. Thus, adding such quantitative cut-offs to study selection criteria may add some precision to the participant screening process for insomnia research studies.

Yet, use of quantitative insomnia criteria based on such subjective sleep estimates may not represent optimal research practice when alternate objective sleep assessment methodologies are readily available. Since polysomnography (PSG) has long been considered the gold standard objective measure of sleep, there has been considerable interest in using quantitative PSG criteria for identifying insomnia research candidates. This has

been particularly true for studies designed to test various pharmacological insomnia therapies. For example, parameters such as sleep onset latency, wake time during sleep, and total sleep time derived from pretreatment nights of PSG monitoring have been commonly used in conjunction with clinical diagnostic assessments to select patients deemed appropriate for medication testing. In such cases, PSG measures most typically have been used to select patients with sufficient levels of sleep disturbance to evaluate the effects of medications with particular pharmacologic characteristics on specific aspects of sleep. This practice seems to fit well with the other mentioned developments focusing on increased precision in insomnia research. Moreover, PSG-based criteria have face validity since they overcome the inaccuracies and reporting biases that can encumber the subjective sleep estimates provided by many insomnia sufferers.

However, the practice of defining and selecting insomnia sufferers for clinical trials on the basis of PSG measures is not devoid of its own shortcomings and potential criticisms. As repeatedly noted in the AASM standards of practice literature[5,6] and elsewhere,[1] only a subset of those who meet diagnostic criteria for insomnia show objective sleep disturbances during PSG monitoring. Also, as noted in the May 2011 proceedings of a FDA-sponsored workshop concerned with the safety and efficacy of insomnia drugs, "…*most of the existing studies are based on one night of polysomnography for sleep…… classification. And we don't really know or haven't systematically examined the reliability of that categorizing criterion*".[7] Although subject selection for many clinical trials has been based on findings from two or more qualifying nights of PSG, the specific quantitative criteria used in these trials have varied considerably. Indeed, multiple distinctive criteria sets have been proposed, and there is no apparent consensus as to which of these is optimal. Furthermore, it remains difficult to evaluate these various criteria since there have been no studies to test the sensitivity and specificity of any PSG measures or criteria sets for discriminating insomnia sufferers from those without insomnia. Thus, whether PSG-based selection criteria add precision to insomnia research remains an open question.

The purpose of this investigation was to test the sensitivity and specificity of PSG measures for the selection of patients with primary insomnia. Specifically, this study tested the usefulness of common sleep measures such as sleep onset time, total sleep time, wake time after sleep onset, and sleep efficiency, derived from lab-based and home-based sleep monitoring for discriminating primary insomnia sufferers from normal sleepers. The sensitivity and specificity of these PSG measures for sample characterization were evaluated by consensus standards[8] for judging their adequacy and were compared to the performances of concurrent measures derived from sleep diaries. Additionally, this study tested the sensitivity and specificity of a number of previously proposed PSG criteria sets for discriminating primary insomnia sufferers from the normal sleepers in our sample.

METHOD

Design

This study comprised a secondary analysis of archival data using a between-groups cross-sectional research design. The study sample included independent groups of age- and gender-matched non-complaining normal sleepers and persons with primary insomnia. The participants for this report were drawn from a larger parent study[9-11] conducted to compare the night-time sleep and daytime functioning of adult insomnia sufferers and normal sleepers. All study procedures were reviewed and approved by the institutional review boards of the VA Medical Center and Duke University Medical Center in Durham, NC. Each participant was required to provide written informed consent prior to enrolling and received a maximum of $250 for completing all procedures of the parent study.

Participants

Both non-complaining normal sleepers and insomnia sufferers were recruited in 3 waves between October 1992 and October 2001 via posted study announcements on bulletin board within the VA and Duke University Medical Centers, letters mailed to persons in the Duke University Center for the Study of Aging and Human Development Subject Pool, and flyers posted in the community. In addition, we recruited a subset of the insomnia sufferers through face-to-face solicitations of clinic patients presenting to our university sleep disorders center. During the first recruitment wave, we enrolled age- and gender-matched insomnia suffers and normal sleepers between 60 and 79 years of age. The second recruitment wave was used to enroll similarly matched middle-aged (i.e., aged 40 to 59 years) insomnia sufferers and normal sleepers, whereas the final recruitment period was used to enroll samples of young adult (aged 20 to 39 years) insomnia sufferers and matched normal sleepers. All study candidates underwent a thorough screening process to ensure that they met study selection criteria. Screening procedures included structured psychiatric (SCID)[12] and sleep interviews[13,14] and a physician-conducted medical exam with thyroid (TSH level) screening. Candidates who passed these initial screening procedures underwent a minimum of 1 to 2 nights of qualifying polysomnography (PSG) conducted in either the sleep lab or in the candidate's home. To be included in the study, insomnia sufferers had to meet DSM-IV criteria for primary insomnia. These criteria include the following: (a) A predominant complaint of difficulty initiating or maintaining sleep or nonrestorative sleep for ≥ 1 month; (b) The sleep disturbance (or associated fatigue) causes clinically significant distress and impairment in social, occupational, or other important areas of functioning; (c) The sleep disturbance does not occur exclusively during the course of another primary sleep disorder; (d) The sleep disturbance does not occur exclusively during the course of another mental disorder; (e) The sleep disturbance is not due to the direct physiological effects of a substance or a general medical condition. We used these criteria for selection of our insomnia sample since they are widely used in clinical venues to identify insomnia sufferers, and currently no biological/medical assay exists for diagnosing insomnia. The normal sleepers selected for this study had to report general satisfaction with sleep in the absence of any reported sleep onset or maintenance difficulties. Exclusion criteria for all participants were: (a) terminal illness; (b) medical condition associated with compromised sleep (e.g., rheumatoid arthritis, thyroid disease); (c) abnormal TSH levels on a screening thyroid panel; (d) history of any prior psychiatric illness (lifelong perspective); (e) current major psychiatric

(Axis I) documented by the SCID; (f) current substance abuse; (g) sedative hypnotic dependence and an unwillingness or inability to abstain from hypnotics while in the study; (h) current use of anxiolytics, antidepressants, or any other psychotropic medication; or (i) an apnea/hypopnea index $\geq$ 15 or a periodic limb movement-related arousal index $\geq$ 15 during screening PSG. Also excluded were insomnia sufferers who met structured sleep interview criteria[13,14] for a comorbid sleep disorder in addition to primary insomnia, as well as normal sleepers who met structured sleep interview criteria for any sleep disorder.

A total of 208 study participants were enrolled; most (> 95%) of these were recruited from posted announcements or solicitation letters. Of the 208 who entered the study, 9 withdrew from the study before undergoing PSG, and 45 were excluded from this investigation because they either failed to complete all scheduled PSG studies described below or because one or more of their PSG recordings were deemed un-scorable due to technical problems (e.g., loss of key electrodes). As a result, 25 (12.5%) participants had 5 nights of usable data, 11 (5.5%) had 4 nights of data, 5 (2.5%) had 3 nights of data, 3 (1.5%) had 2 nights of data, and 1 had only one 1 night of usable data (0.5%). Of the remaining sample, 76 (38 women) met selection criteria for primary insomnia, and the remaining 78 (38 women) were non-complaining normal sleepers. The insomnia sample reported a 10.6-year (SD 9.3 years) history of sleep difficulties on average and comprised 11 persons with solely sleep onset complaints, 25 individuals with solely sleep maintenance complaints (either middle-of-the-night wakefulness and/or early morning awakenings), 38 with a mixture of sleep onset and sleep maintenance difficulties, and 4 persons with other sleep/wake concerns (e.g., non-restorative sleep). The mean age of the insomnia sample was 51.8 years (SD 16.5 years), whereas the mean age of the normal group was 50.9 years (SD 16.0 years). The insomnia group included 59 Caucasians, 10 African Americans, 5 Asian Americans, 1 Hispanic American, and 1 Native American; the normal sleepers included 67 Caucasians, 9 African Americans, 1 Asian American, and 1 Native American.

Polysomnography

As part of their requirements for the parent study, all participants were scheduled for 3 consecutive nights of polysomnography (PSG) conducted in their homes and an additional 3 consecutive nights of PSG in our university medical center's sleep laboratory. PSGs were conducted using 8-channel Oxford Medilog 9000 or 9200 series ambulatory recorders. The monitoring montage included 2 electroencephalogram (EEG) channels (C_3-A_2, O_z-C_z), bilateral electrooculogram (EOG), submental electromyogram (EMG), 2 channels of anterior tibialis EMG (right and left leg), and a nasal-oral thermistor. All PSGs were scored using traditional scoring criteria for assignment of sleep stages, identification of respiratory events (e.g., apneas, hypopneas), and identification of periodic limb movements and periodic limb movement-related arousals.[15-18] Per pre-planned study protocols, the first or initial 2 PSG nights (home or lab) were used to screen out those exceeding the aforementioned apnea-hypopnea index or periodic limb movement arousal index cutoffs for study inclusion. Although PSG typically includes additional respiratory measures to detect breathing abnormalities (e.g., respiratory effort, oximetry), it was thought that monitoring of na-

sal/oral airflow along with our thorough interview screening for apnea would be sufficient to identify most cases with an apnea-hypopnea index above the study's exclusionary cut-off.

Laboratory personnel who were kept blind to the participants' diagnoses (insomnia vs. normal sleeper) scored all PSG recordings. Each taped PSG record was scored directly on the screen of the Medilog playback unit using traditional scoring criteria.[16] Results of PSG scoring were subsequently used to derive measures of time in bed (TIB: time between the electronically marked bedtime and final rising time on each recording), total sleep time (TST: the total time in stages 1, 2, slow wave sleep, and REM sleep), latency to persistent sleep (LPS: time between lights out and the first 10 min of sleep containing $\leq$ 2 min of wake time, stage 1 sleep, or movement time), wake time after sleep onset, (WASO: all time awake after the onset of sleep and before the final morning rising time), and sleep efficiency % (SE %; [TST $\div$ TIB] $\times$ 100%). It should be noted that the sleep onset measure chosen for use here—LPS—differs from other measures of sleep onset that require a more limited number of epochs or shorter time periods of sleep for connoting the transition from wakefulness to sleep. We chose LPS since it has been used in many clinical trials as a measure of sleep onset and because it connotes not only the ability to achieve sleep but also to sustain it. Hence, it is a particularly relevant parameter for those with sleep onset complaints.

Sleep Diary Monitoring

In addition to PSG recording, participants were asked to complete sleep diary forms in the morning after each PSG night and during an additional 2-week period when they were not undergoing PSG monitoring. The diary forms asked participants about bed and rising times as well as their how long it took them to fall asleep, how long they were awake in the middle of the night after first falling asleep, and how long they were awake in bed at the end of the night before getting out of bed. The diary forms additionally included items asking participants to rate the quality of each night's sleep (1 = very poor; 5 = excellent) and how well rested they felt upon arising (1 = not at all; 5 = very well rested). These items were common to both sets of diaries. In addition, the diary forms completed by participants after each PSG night included an item that asked how much the PSG recording equipment disturbed their sleep each night (1 = very much; 5 = not at all). The diaries completed after each PSG served mainly to help PSG scorers corroborate participants' bed and rising times as well as to assess the degree to which the sleep monitoring process was perceived to disturb participants' sleep. In contrast, the 2-week diaries were used to obtain measures of participants' subjective measures of time in bed, sleep onset latency (SOL), middle of the night wake time (MWASO), wake time at the end of the night (TWASO), total sleep time (TST), and sleep efficiency (SE = % of time in bed spent asleep). These measures were calculated for each night of sleep recorded by participants on the 2-week diaries and then used to calculate mean values across the 2 weeks of diary monitoring.

Procedure

Consenting participants who met selection criteria underwent PSGs in a randomly determined order so that roughly half of the participants in each group (normals and insomnia) underwent

Table 1—Comparisons of mean PSG values from lab and home monitoring

| | Lab PSG Comparisons | | | | | | Home PSG Comparison | | | | | |
| | Normal | | Insomnia | | | | Normal | | Insomnia | | | |
	M	SD	M	SD	F	P	M	SD	M	SD	F	P
LPS	26.6	63.3	23.5	19.1	0.21	0.65	16.5	10.5	39.0	125.5	3.09	0.08
WASO	48.4	36.0	62.2	33.8	7.4	< 0.01	56.3	49.7	69.8	48.3	3.7	0.06
TST	384.1	53.8	378.1	2.6	0.52	0.47	386.3	57.1	378.1	68.8	0.81	0.37
SE	88.0	8.9	85.1	9.2	5.1	0.03	86.5	9.4	83.1	10.2	5.8	0.02
TIB	442.3	66.5	450.7	50.5	0.98	0.32	450.5	57.7	458.4	52.0	0.99	0.32
STAGE1	17.6	10.0	19.0	10.8	0.90	0.34	19.4	10.6	19.3	13.4	< 0.01	0.97
STAGE2	196.1	37.2	194.0	40.3	0.15	0.70	198.2	37.6	191.7	42.6	1.3	0.26
SWS	79.7	30.3	77.1	28.1	0.39	0.54	76.9	24.8	78.2	26.0	0.13	0.72
REM	80.2	20.2	78.5	24.0	0.27	0.61	82.5	21.2	78.1	21.1	2.1	0.15
EPI	81.3	31.0	94.4	33.3	8.06	< 0.01	86.4	33.1	97.2	41.6	4.1	0.05

EPI, the number of sleep episodes, with episode defined as one or more continuous epochs of sleep preceded and followed by wakefulness; LPS, latency to persistent sleep in minutes; REM, rapid eye movement sleep; SE, sleep efficiency percent; STAGE1, stage 1 sleep; STAGE2, stage 2 sleep; SWS, slow wave sleep; TIB, time in bed; TST, total sleep time in minutes; and WASO, wake after sleep onset in minutes.

lab PSG, first whereas the other half underwent home studies first. All sleep studies were scheduled such that the home and laboratory PSGs were separated by ≥ 4 but ≤ 30 intervening days. During both series of studies, participants were instructed to maintain customary home bed and rising times, and they were instructed to note their actual bed and rising times each night using an event marker contained on the PSG recorders. They also recorded bed and rising times on a sleep diary they completed upon arising each morning after each PSG night. In addition, all home studies were scheduled for nights when participants planned to have no overnight houseguests. Participants were instructed to abstain from alcoholic beverages and to not consume caffeinated substances after 18:00 on study nights.

RESULTS

Preliminary Analyses

Before conducting our primary analyses, we computed the means and standard deviations of each of the 4 PSG sleep measures for each sample in each recording site. **Table 1** shows these data along with results of a series of one-way ANOVAs computed to compare the normal sleepers and insomnia sufferers on these measures in each setting. These analyses showed the insomnia sufferers had significantly higher mean WASO and a significantly lower mean sleep efficiency in the lab setting than did the normal sleepers. In the home setting, the groups differed only on their sleep onset latencies.

We also conducted preliminary analyses to determine if those excluded from the final sample (due to their having < 6 nights of PSG data) differed from those who comprised the final sample used to address study objectives. Specifically, we compared these groups in terms of their gender composition, mean age, and mean values of PSG sleep measures obtained in the home and sleep lab settings. A χ^2 analysis showed the group of excluded participants was not significantly different from the sample retained for study analyses in terms of their gender composition (p > 0.05). However, a one-way ANOVA showed

the group of excluded participants was significantly younger (34.3 years) than those retained (51.3 years) in the final sample (F = 38.89; p = 0.0001). This latter finding was attributable to the fact that the youngest cohort was recruited last and at a time when the recording equipment had undergone extensive use and was more prone to malfunction and fall into disrepair. Consequently, more nights of data loss occurred with the youngest cohort than with the older groups enrolled.

Linear mixed models (LMM) were conducted using the Proc Mixed procedure included in the Statistical Analysis System software version 9.2 to compare our final sample and excluded participants on each of the 10 sleep measures derived from PSG recordings. LMM were chosen since they allow inclusion of all cases including those with missing nights of data in the planned comparisons. Separate series of analyses were conducted with the sets of measures derived from home and lab PSGs, and the statistical model included subject *type* (insomnia vs. normal sleeper), *group* (included vs. excluded participants) and *nights* as independent factors. Because the excluded group was significantly younger than the sample retained for study analyses, age was included as a covariate in the series of analyses conducted. Since a total of 20 LMM analyses were conducted, a Bonferroni-corrected α = 0.0025 (i.e., .05 ÷ 20) might be considered for assigning statistical significance. Whereas this criterion level of significance is appropriate for main study analyses, it arguably is a high bar to reach in our tests of possible study confounding. Hence, we used a more liberal α = 0.01 level for assigning statistical significance in these 20 LMM analyses. Despite selection of this lenient α level, results of all these analyses showed no significant main or interaction effect for the group factor. Hence, the sleep measures derived from the inclusion sample were representative of the larger participant group enrolled.

ROC Analyses

In order to ascertain how well our selected PSG measures discriminated insomnia sufferers from normal sleepers, we first compared these groups in regard to their means values of the home and lab PSG measures obtained. **Table 1** shows these mean

values and results of statistical comparisons. These results show the insomnia sufferers had more WASO, lower sleep efficiencies, and a great number of sleep episodes in the lab than did normal sleepers, whereas in the home setting, the insomnia group showed higher WASO values and more sleep episodes than did the normal group. As a follow-up to these simple mean comparisons we used receiver-operating characteristic (ROC) curve analyses to graphically depict the relation between the sensitivity and specificity of the test over all possible values of each mean PSG measure. In these analyses, sensitivity for a particular mean PSG measure represented the probability of detecting insomnia when it is present, and specificity represented the probability of not detecting insomnia when it was not present. The ROC curve was plotted for all values, and the greater the distance of ROC curve above the diagonal reference line, the more accurate the test.[19] The area under the curve (AUC) served as one primary index of accuracy. The AUC is the probability that a test result for a randomly chosen positive case will exceed the result for a negative case. Swets[8] has suggested that test accuracy be defined as "low" for AUC values under 0.7, "moderate" for AUC values between 0.7 and 0.9, and "high" for values greater than 0.9. In addition, we calculated the Youden Index for each of the sleep measures tested. The Youden Index is an alternative index of diagnostic accuracy. It ranges from 0 to 1, with higher values representing greater separation between 2 distributions. Thus, distributions having complete overlap would have a Youden Index of 0, while those having complete separation would have a Youden Index of 1.

In conducting the ROC analyses, we focused initially on the 4 variables most reflective of an insomnia diagnosis: difficulty falling asleep (LPS); difficulty staying asleep (WASO); sleep duration (TST); and sleep efficiency (SE). Results of the ROC analyses conducted with these home- and laboratory-based measures are summarized in **Figures 1** and **2**, respectively. The top portions of **Table 2** provide summary statistics for these analyses including AUC values, Youden indices, optimal cutoffs for group discrimination, and the sensitivity and specificity of these cutoffs. Collectively, the figures and table show that none of the home or lab PSG measures performed well in discriminating the insomnia sufferers from the non-complaining normal sleepers. The best discriminators among the home-based measures were SE and LPS, whereas the best discriminators among those measures derived from lab recordings were SE and WASO. Nonetheless, AUC estimates all fell in the "low" range, indicating the poor test accuracy of these measures for use as quantitative insomnia criteria. The inadequacy of these mean values for discriminating the insomnia and normal groups is characterized by the plots of data shown in **Figure 3**. This figure shows the distribution of mean values of TST and SE obtained from lab recordings. These plots show that there is appreciable overlap between the insomnia and normal groups in regard to the distributions of measures across the ranges of values observed. Similar plots (not shown) for all of the remaining lab and home sleep measures included in **Table 1** showed comparable results. As a consequence, it is difficult to discern a single quantitative cutoff for any of these mean values, even for measures showing significant group differences by the ANOVA tests.

Since an insomnia diagnosis does not require the presence of both a sleep onset and sleep maintenance complaint, we also explored the possibility that PSG-based sleep measures may better discriminate if the type of insomnia complaint (onset versus maintenance) is used as the comparator. We conducted 2 additional analyses comparing normal sleepers to those with a sleep onset complaint and to those with a sleep maintenance complaint. For the purpose of these analyses, the sleep onset group comprised the insomnia sufferers with sleep onset only or mixed onset/maintenance complaints (total = 49), whereas the maintenance group included those with maintenance complaints with or without accompanying sleep onset concerns (total = 63). Results of these tests (see supplemental material) showed that all AUC values fell in the low range, suggesting poor test accuracy. Thus, considering the type of insomnia complaint did not increase the accuracy of PSG sleep variables for discriminating insomnia sufferers from normal sleepers.

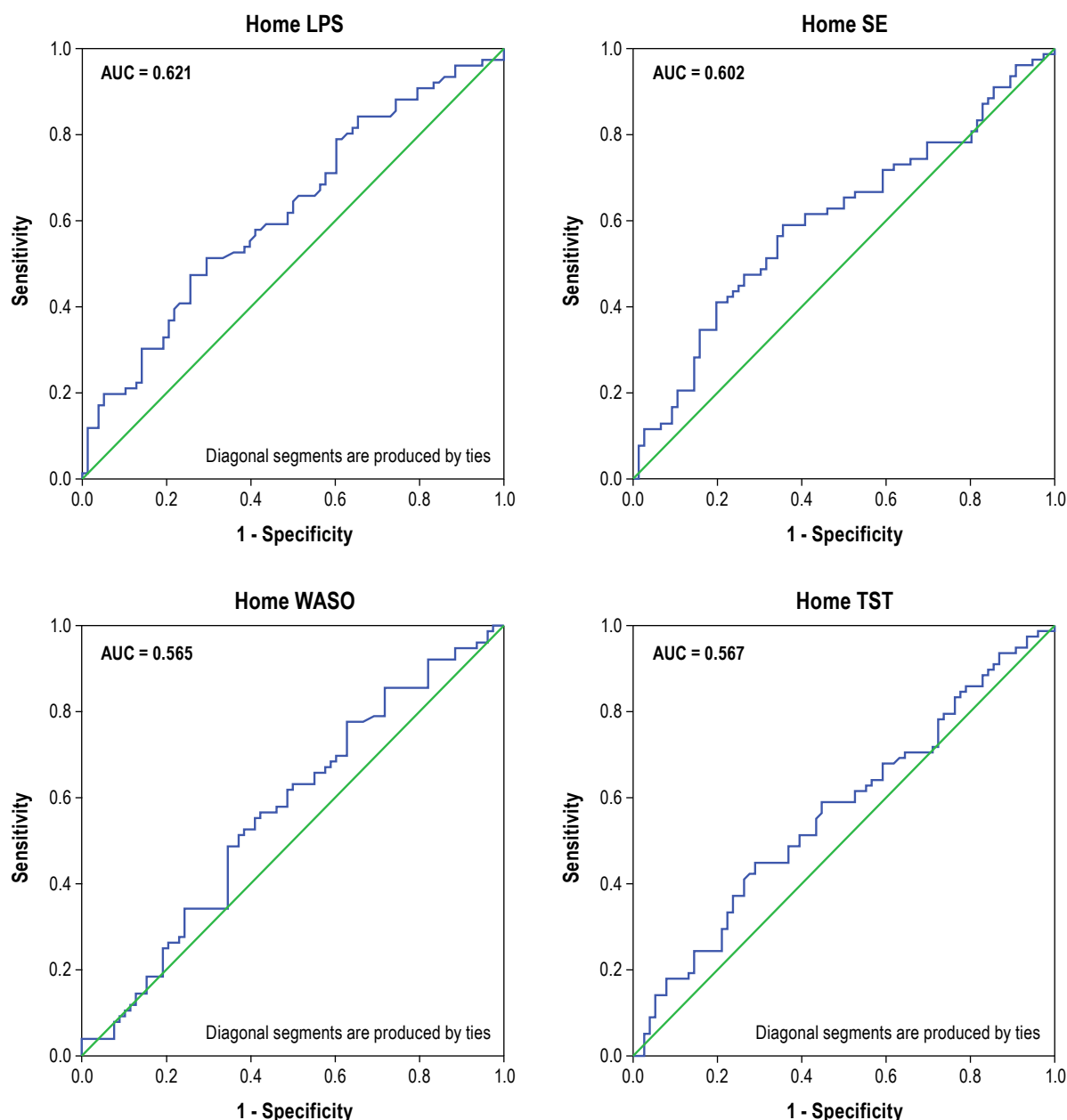## Relative Sensitivity and Specificity of Sleep Diary Measures

Since measures derived from extended periods of sleep diary monitoring have demonstrated sensitivity/specificity for defining insomnia[3,4] we conducted ROC analyses using 2-week sleep diary data collected by participants during a period when they were not undergoing PSG. The bottom portion of **Table 2** shows the AUCs and Youden indices derived from ROC analyses of these sleep diary measures. Also shown are the optimal cutoffs for group discrimination as well as the sensitivity and specificity of those cutoffs. The sleep diary measures were found to be consistently more accurate than PSG in discriminating insomnia sufferers from normal sleepers, particularly the measures of SE, SOL, and middle of the night wake time (MWASO).

Tests of Previously Reported PSG Qualifying Criteria

It is possible that mean values of individual sleep measures provide a somewhat limited view of the detectable PSG differences between insomnia sufferers and those without sleep complaints. Indeed, it may be useful to consider indicators of the persistence of sleep difficulties among those with insomnia relative to normal sleepers, and/or to simultaneously consider multiple PSG measures to enhance group discrimination. In this regard, several recently published studies conducted to test hypnotic agents have described sets of PSG criteria used in conjunction with clinical diagnoses to select primary insomnia patients as study participants. The specific criteria sets suggested in these trials are shown in **Table 3**.[20-28] As noted, most of these criteria sets considered both average values of selected sleep measures across 2 qualifying nights of PSG and the magnitude of these measures on each night separately. As such, these selection criteria not only reflect a specific level of sleep difficulty on average, but also a degree of persistence in that sort of disturbance. The criteria proposed by Mayer[27] would seem designed to identify those with sleep onset difficulties, whereas those provided by Roth[28] appear suited to select patients with sleep maintenance complaints. The remaining criteria sets shown would appear useful for identifying insomnia patients with a mixture of sleep onset and maintenance difficulties.

In selecting these various criteria sets, we recognized that they were not intended to be used in isolation to define insomnia, but rather were employed in conjunction with clinical assessments to identify individuals with a particular form or severity of sleep disturbance. We also recognized that these criteria sets

**Figure 1**—ROC plots for home PSG measures

Home LPS

AUC = 0.621

Diagonal segments are produced by ties

Sensitivity vs 1 - Specificity

Home SE

AUC = 0.602

Sensitivity vs 1 - Specificity

Home WASO

AUC = 0.565

Diagonal segments are produced by ties

Sensitivity vs 1 - Specificity

Home TST

AUC = 0.567

Diagonal segments are produced by ties

Sensitivity vs 1 - Specificity

were originally intended for PSGs conducted with participants who were provided prescribed (fixed) amounts of time in bed rather than the participant-determined bed and rising times used in this study. Nonetheless, since these criteria sets by and large considered multiple sleep parameters across nights, we reasoned that they might provide better discrimination of our insomnia and normal sleeper groups than would individual measures tested in our above-described ROC analyses. To test this assumption, we conducted a series of cross tabs analyses to evaluate the accuracy of each of these criteria sets for discriminating our insomnia group from our normal sleeper group. In doing so, we selected relevant measures from the first 2 nights of PSG recording conducted in the laboratory and the initial 2 nights obtained in participants' homes. The classification rules were then applied to the lab- and home-derived data separately so as to ascertain

how the sleep setting effects influenced our classification results. We used our cross tabs analyses specifically to ascertain each criteria set's sensitivity (i.e., the probability for detecting insomnia when it is present), and specificity (i.e., the probability of not detecting insomnia when it was not present).

Results of these cross-tabs analyses are presented graphically in **Figures 4** and **5**. **Figure 4** depicts results for the various criteria sets when applied to lab-based PSG parameters, whereas **Figure 5** shows results obtained with the home-based PSG measures. These figures demonstrate that none of the criteria sets showed good sensitivity and specificity in discriminating our insomnia and normal sleeper groups. Furthermore, the recording site—lab or home—from which PSG data were derived seemingly had little effect on classification results. The majority of the criteria sets showed satisfactory to good speci-

Figure 2—ROC plots for lab PSG measures
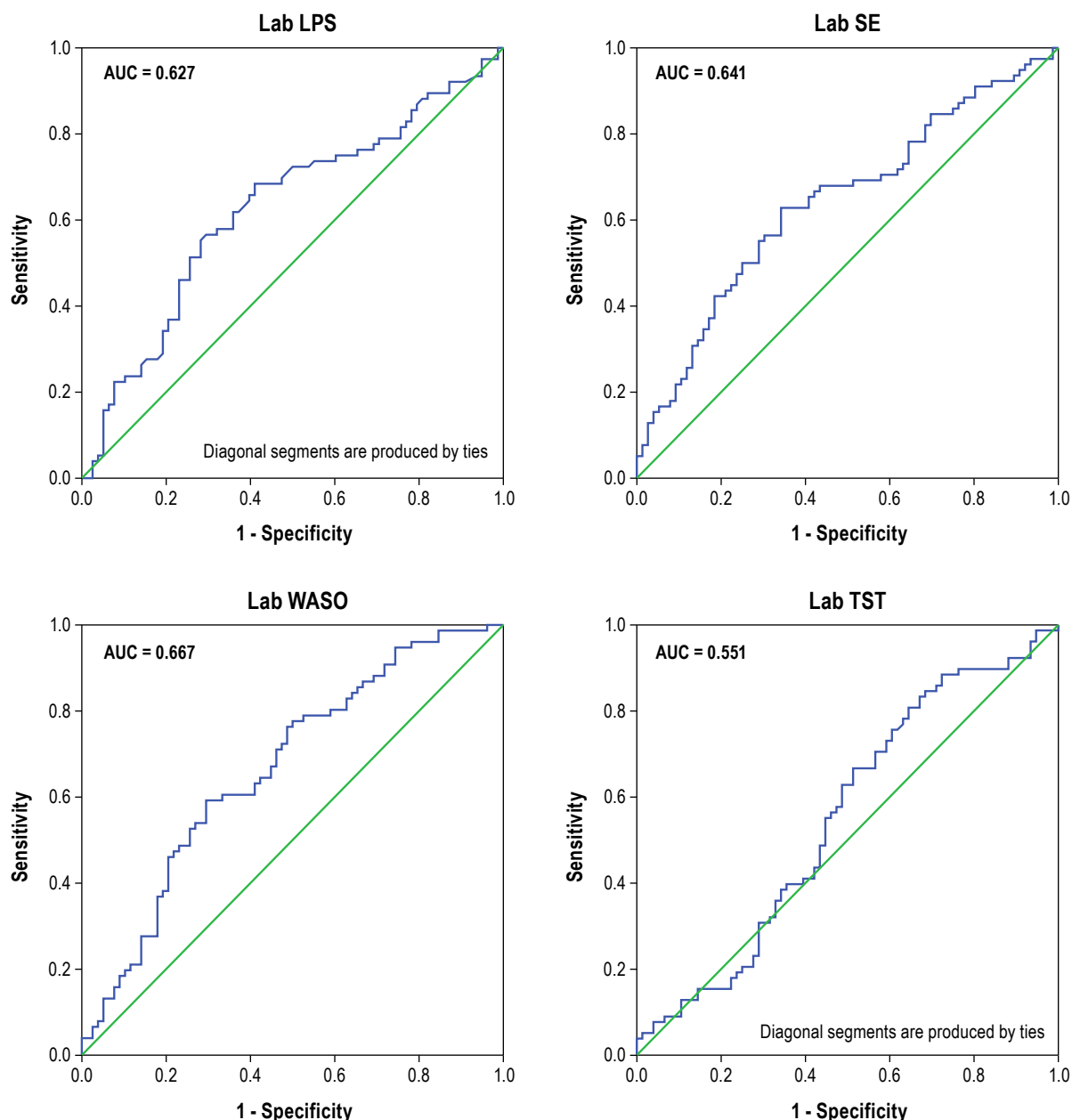
ficity but poor sensitivity for identifying our insomnia sufferers. In fact, less than a third of the insomnia sufferers met most of these selection criteria. The Walsh et al. PSG criteria,[25] based on TST and MSLT sleep latency measures, had relatively poor sensitivity and specificity. Hence, none of the criterion sets seem to convey sleep characteristics that are specific to insomnia.

In addition to testing each criterion set as proposed, we considered the possibility that some components of these various criteria sets may be effective for discriminating the insomnia and normal sleeper groups. As the LPS criterion used across trials was standard, we chose to test this criterion in isolation and in combination with WASO and/or TST criteria. For the purpose of these analyses we conducted cross tabulations of each of the WASO and TST criteria shown in **Table 3** to ascertain the best for group discrimination. Results of these analy-

ses showed that the WASO criterion used by Krystal et al.[26] and the TST criterion proposed by Roth et al.[28] produced the best group separation. We then conducted cross tabulations of each of these criteria individually and in various combinations to ascertain their sensitivity and specificity for defining our insomnia group. Results of these analyses are shown in **Table 4**. These data show that none of the individual criteria or their various combinations had acceptable sensitivity and specificity for identifying our insomnia cohort. Thus, our efforts to optimize these published criteria sets proved unsuccessful.

DISCUSSION

The current study was conducted to test the usefulness of quantitative PSG criteria for the identification/selection of pri-

Table 2—Comparison of ROC AUCs, optimal cutoff values and sensitivity/specificity of PSG and diary measures for discriminating those with and without insomnia

Sleep Measure	AUC	Youden Index	Cutoff Value	Sensitivity	Specificity
Home PSG					
TST	0.567	0.09	374.78	0.564	0.553
SE	0.602	0.12	84.28	0.590	0.592
LPS	0.621	0.21	17.36	0.579	0.577
MWASO	0.565	0.06	48.33	0.566	0.564
TWASO	0.521	0.04	4.05	0.513	0.526
Lab PSG					
TST	0.551	0.07	376.07	0.551	0.553
SE	0.641	0.18	86.47	0.628	0.632
LPS	0.627	0.05	16.14	0.618	0.628
MWASO	0.667	0.21	45.20	0.605	0.603
TWASO	0.606	0.02	3.40	0.553	0.551
2-Week Sleep Diary					
TST	0.715	0.41	408.32	0.659	0.657
SE	0.905	0.66	86.70	0.817	0.814
SOL	0.888	0.63	16.96	0.841	0.843
MWASO	0.851	0.56	15.18	0.756	0.743
TWASO	0.675	0.23	19.82	0.610	0.614

TST, total sleep time; SE, sleep efficiency; LPS, latency to persistent sleep; SOL, sleep onset latency; MWASO, middle WASO; TWASO, terminal wake time.

Figure 3—Distribution of mean PSG values in the two samples

mary insomnia sufferers in research protocols. The availability of such criteria derived from objective sleep monitoring would be extremely useful for standardizing the samples used in insomnia research and thereby facilitate comparisons of results across insomnia research studies. Unfortunately, the analyses conducted herein to identify useful quantitative PSG criteria did not support any of the criteria sets examined. Our ROC analyses, for example, showed that mean values of LPS, WASO, TST, and SE, derived from series of lab or home monitoring, failed to accurately discriminate primary insomnia sufferers from normal sleepers. Our analyses also showed that none of the PSG-based insomnia selection criteria sets used in recent insomnia treatment trials demonstrated acceptable sensitivity and specificity for selecting those meeting diagnostic criteria for primary insomnia. In fact, all of these criterion sets showed a tendency to select 50% or fewer of the insomnia sufferers and, for most such sets, our normal sleepers met these criteria at roughly the same rates as did the insomnia sufferers. Moreover, none of the cutoffs included in these criteria sets considered individually or in various combinations accurately selected the insomnia cases. To ensure that we did not overlook other sleep architecture variables which might have performed better in terms of identifying those with insomnia, we conducted additional ROC analyses on other PSG variables including the total number of sleep episodes recorded as well as the times spent in stage 1, stage 2, REM, and slow wave sleep. We found low test accuracy across all of these PSG variables as well. See the supplemental material for the specific results of these analyses.

These findings contrast markedly from those reported for measures taken from sleep diaries. Lichstein et al.,[3] for example,
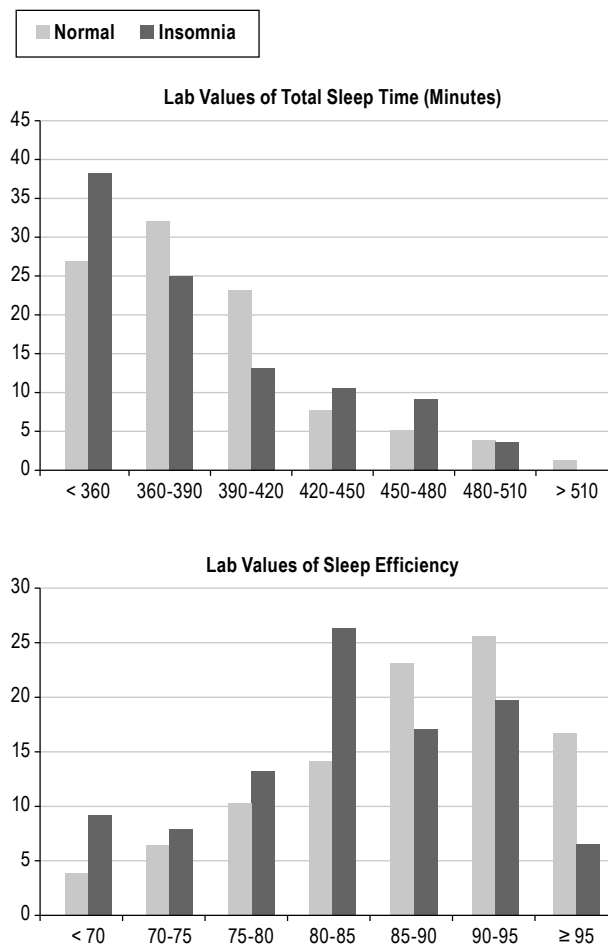
found that sleep diary measures of sleep onset latency or wake time after sleep onset ≥ 31 minutes occurring ≥ 3 times per week have good sensitivity and specificity for discriminating insomnia sufferers from normal sleepers. Likewise, we previously were able to identify specific mean values of sleep onset latency and WASO taken from two weeks of sleep diary monitoring that effectively separate primary insomnia sufferers from normal sleepers in younger and older samples.[4] In the current study, we expanded on our previous diary findings by showing that individual mean diary values of SOL, SE, and MWASO taken from two weeks of conventional home diary monitoring far outperformed any of our PSG measures for group discrimination. The more limited period of PSG monitoring may have put the PSG measures at a relative disadvantage and, in part, accounted for their poorer performance herein. In addition, both diary data and a clinical insomnia diagnosis are based on self-report and therefore share method variance not shared by objective PSG monitoring. Both these methodological factors should be recognized when considering our findings.

Perhaps our results are not surprising inasmuch as the overlap between the PSG measures derived from insomnia sufferers and normal sleepers has long been recognized. Indeed, there is no established and commonly recognized "finding" on the polysomnogram that is accepted as proof of an insomnia diag-

Table 3—Published PSG qualifying criteria used for selection of primary insomnia patients

#	Authors	PSG Criteria Sets
1	Zammit et al. (2004)	2 nights of PSG with: Mean LPS ≥ 20 min with no night < 15 min Mean TST ≤ 420 min or mean WASO ≥ 20 min with neither night < 15 min
2	McCall et al. (2006)	2 nights of PSG with: Mean WASO ≥ 20 min with each night < 15 min Mean LPS ≥ 20 min with neither night < 15 min
3	Roth et al. (2007) Scarf et al. (2008)	2 nights of PSG with: LPS ≥ 10 min Mean wake time during sleep (WTDS) ≥ 60 min with each night < 45 min 240 min < TST ≤ 410 min
4	Lankford et al. (2008)	2 nights of PSG with: Mean WASO ≥ 45 min & WASO > 30 each night Mean latency to persistent sleep (LPS) > 20 min & LPS > 15 min each night Mean TST ≤ 6.5 hours & TST < 7 hours each night
5	Walsh et al. (2010)*	2 nights of PSG with: 240 min < Mean TST ≤ 410 min 4 min ≤ Mean MSLT Latency ≤ 16 min
6	Krystal et al. (2010)	2 nights of PSG with: LPS > 10 min WTDS ≥ 60 min 240 min < TST ≤ 390 min
7	Mayer et al. (2009)	2 nights of PSG with: Mean LPS > 20 min & LPS ≥ 15 min each night
8	Roth et al. (2006)	2 nights of PSG with: Mean WASO > 40 with each night ≥ 30 min 180 min ≤ TST ≤ 420 min each night

*The Walsh et al. criteria also included cutoffs for the average sleep latency on the multiple sleep latency test, but only the nocturnal PSG criteria were examined herein.

Table 4—Additional tests of the sensitivity and specificity of various PSG criterion sets considered individually and in combination

Criteria Used	Sensitivity % Insomnia Identified	Specificity % Normals Identified	χ^2 (p)
Home			
LPS	36.5	80.4	6.82 (< 0.01)
WASO	52.1	63.9	5.01 (0.03)
TST	74.0	26.8	0.014 (0.90)
LPS + WASO	17.7	89.7	2.20 (0.14)
LPS + TST	29.2	84.5	5.23 (0.02)
WASO + TST	40.6	70.1	2.43 (0.12)
LPS + WASO + TST	13.5	91.8	1.40 (0.24)
Lab			
LPS	29.5	75.8	0.67 (0.41)
WASO	46.3	72.6	7.33 (< 0.01)
TST	73.7	18.9	1.47 (0.23)
LPS + WASO	14.4	91.8	1.92 (0.17)
LPS + TST	21.6	80.6	0.15 (0.70)
WASO + TST	39.2	76.5	5.59 (0.02)
LPS + WASO + TST	12.6	92.6	1.46 (0.23)

The LPS criterion tested was an LPS ≥20 min average across nights, with neither night having <15 min; the WASO criterion tested was a wake time during sleep ≥60 min each of 2 nights; the TST criterion tested was total sleep time ≥180 min but ≤420 min on each night.

nosis. In the clinical setting, it is common experience, in fact, to find relatively normal sleep among some insomnia patients undergoing PSG recording, although such findings may be attributed to the so-called "reverse night effect" or the possible presence of paradoxical insomnia. It is also recognized that PSG recording can be disruptive to normal sleepers and produce sleep results that overlap with the sleep disruption characterizing insomnia samples. Such factors may account for the minimal differences often found in PSG comparisons of insomnia and normal groups. In line with these observations, current practice parameters indicate that polysomnography is generally not required for the routine diagnosis of insomnia, and there are relatively limited indications for its use among patients with insomnia complaints.[5,6] Although the studies referenced herein did not necessarily use PSG for purposes of identifying those with insomnia, it would have been useful to have found that these criteria were helpful for this purpose. However, consistent with the current practice parameters, our study findings suggest that relying mainly on quantitative PSG-based criteria for the selection of study participants in insomnia treatment research is not advisable.

In considering our results it should be noted that the effectiveness found for any test is dependent upon the "gold standard" against which it is judged. Insomnia is defined on the basis of a patient's clinical complaints, and there currently exists no biological assay or medical test that reliably confirms an insomnia disorder. For the purposes of this study, we chose to use DSM-IV-TR criteria for primary insomnia in identifying and selecting our insomnia sample; yet the primary insomnia diagnosis is based solely on self-report and is widely accepted to include a somewhat heterogeneous patient group. Moreover, the reliability of this diagnosis across raters seems marginal at best.[30,31] Reliance on the clinical insomnia diagnosis as the gold standard has limitations in its own right and likely contributes to the findings obtained herein. Indeed, it could be argued that the clinical insomnia diagnosis has poor sensitivity and specificity for identifying individuals with large amounts of WASO or prolonged sleep latency on PSG. Nonetheless, at this juncture, the clinical diagnosis remains the defining feature of insomnia. Thus, the use of clinical criteria for establishing the diagnosis seemed reasonable for the purposes of this study.

Of course, the role of PSG selection criteria may go beyond the purpose of mere identification of those who meet diagnostic criteria for a specified disorder. There is also the consideration of selecting study participants who have a sufficient level or severity of disease so as to assure that treatment effects can be clearly discerned and detected by whatever statistical methods are deemed appropriate for the study in question. In the case of insomnia, it is common practice in this regard to select patients with a certain amount of wakefulness as measured by sleep onset latency and/or wake time during the night. This practice has indeed been common in studies of both pharmacological and

Figure 4—Tests of published criteria sets applied to lab PSG values
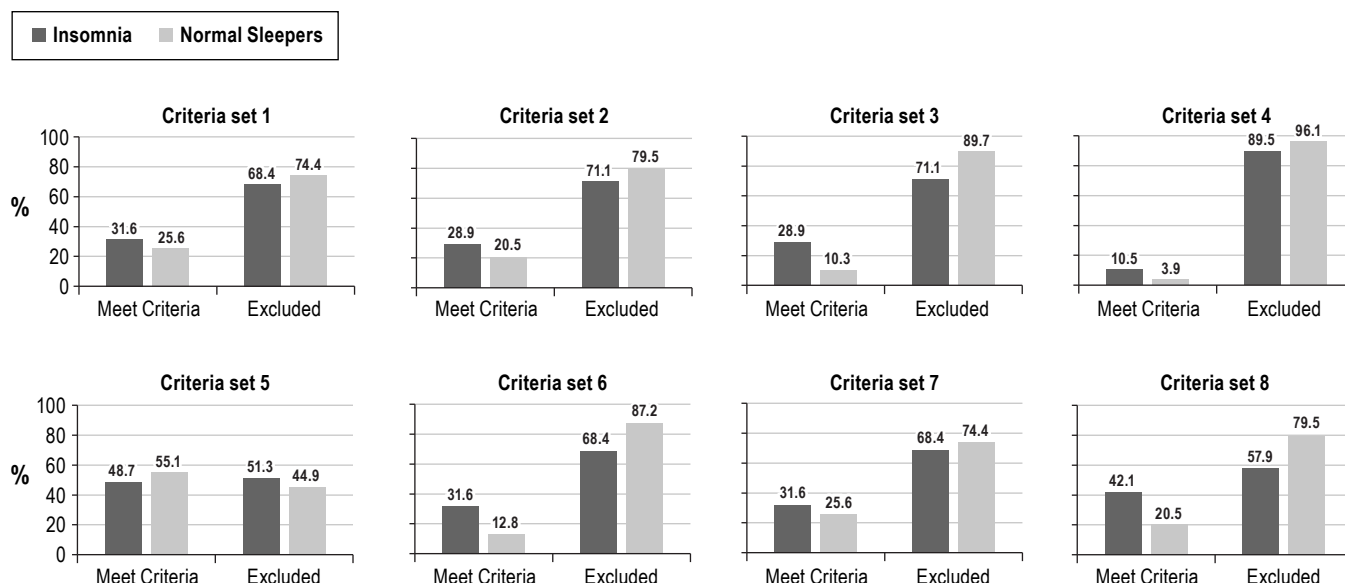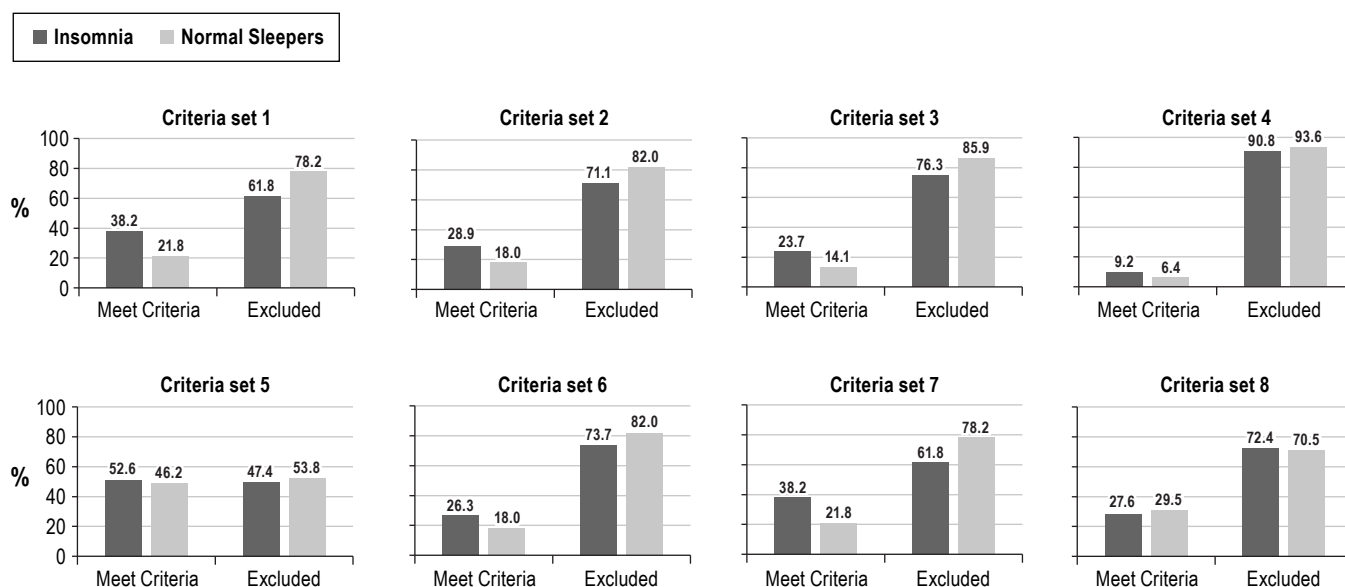

Figure 5—Tests of published criteria sets applied to home PSG values

psychological therapies. As the Food and Drug Administration has required the inclusion of PSG outcome measures in tests of new insomnia agents, it is not surprising that the use of quantitative PSG selection criteria have been commonplace in insomnia medication trials. Yet the findings concerning the PSG selection criteria sets listed in **Table 3** showed a sizable proportion of insomnia patients fail to meet these criteria and thus would be excluded from such studies. This was the case even when we attempted to increase test precision by considering sleep onset and sleep maintenance insomnia subtypes separately. Such results, in turn, raise questions about how well findings for such trials generalize to the larger insomnia population. It is clear from the previously cited proceedings of the May 2011 FDA/PERI sponsored workshop[7] that the FDA is now reconsidering

the role and use of PSG in insomnia clinical trials, so perhaps the sleep research community should do so as well. Of course, PSG may remain a useful research tool for documenting objective sleep changes in insomnia treatment studies. Indeed, as PSG remains the gold standard for sleep measurement, this procedure is perhaps the best suited one for measuring the impact of a treatment on the overall sleep process.

Admittedly, this investigation had some limitations that need to be considered. Although our sample was moderate in size, it was heterogeneous in terms of insomnia complaint (e.g., onset, maintenance, mixed) and was composed largely of research volunteers. The findings therefore may not generalize to clinical populations since research and clinical samples may differ markedly. Furthermore, the sample included mainly Cauca-

sians, so it is not known how representative the findings would be for other ethnic groups. We also should note that many of the criteria sets used latency to persistent sleep (LPS) as a consideration, and the definition of this parameter is typically considered the onset of a 10-minute period of uninterrupted sleep. The definition used for LPS in constructing our archival data set was slightly less demanding. Although our measure of LPS would be expected to correlate highly with the LPS measure use in pharmacology trials, the two measures are admittedly not identical. As a result, it is possible that our analyses may underestimate the performances of at least some of the criteria sets we tested. Yet most of these sets performed so poorly that it is reasonable to assume that the use of more rigorous LPS measures would likely not result in significant performance improvement. Finally, we should acknowledge that the PSG montage used provided inadequate screening for sleep disordered breathing, given the relationship between subtle respiratory disturbance and insomnia complaints, especially in women. Despite this and the other limitations noted, our findings suggest that the use of quantitative PSG-based selection criteria in insomnia research is a practice that should be questioned. Use of such selection criteria in isolation may set an unreasonable metric for patients to achieve, and thus, may exclude many with genuine insomnia disorders from empirical scrutiny.

REFERENCES

1. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. *Sleep* 2006;29:1155-73.
2. Edinger JD, Bonnet M, Bootzin RR, et al. Derivation of research diagnostic criteria for insomnia: Report on an American Academy of Sleep Medicine work group. *Sleep* 2004;27:1567-96.
3. Lichstein K, Durrence H, Taylor D, Bush A, Riedel B. Quantitative criteria for insomnia. *Behav Res Ther* 2003;41:427-45.
4. Lineberger M, Carney C, Edinger J, Means M. Defining insomnia: quantitative criteria for insomnia severity and frequency. *Sleep* 2006;29:479-85.
5. Chesson A Jr, Hartse K, Anderson WM, et al. Practice parameters for the evaluation of chronic insomnia. An American Academy of Sleep Medicine report. Standards of Practice Committee of the American Academy of Sleep Medicine. *Sleep* 2000;23:237-41.
6. Littner M, Hirshkowitz M, Kramer M, et al. Practice parameters for using polysomnography to evaluate insomnia: an update. *Sleep* 2003;26:754-60.
7. Buysse DJ. Defining insomnia: major subtypes, key symptoms. Paper presented at: Safety And Efficacy Of Hypnotic Drugs: An FDA-PERI Co-Sponsored Workshop; May 10 & 11, 2011; Bethesda, MD.
8. Swets J. Measuring the accuracy of diagnostic systems. *Science* 1988;240:1285-93.
9. Edinger JD, Fins AI, Sullivan RJ, et al. Sleep in the laboratory and sleep at home II: Comparison of older insomniacs and normal sleepers. *Sleep* 1997;20:1119-26.
10. Edinger JD, Glenn DM, Bastian LA, et al. Sleep in the laboratory and sleep at home II: Comparison of middle-aged insomnia sufferers and normal sleepers. *Sleep* 2001;24:761-70.
11. Edinger JD, Means MK, Carney CE, Krystal AD. Psychomotor performance deficits and their relation to prior nights' sleep among individuals with primary insomnia. *Sleep* 2008;31:599-607.
12. Spitzer RL, Williams JBW, Gibbons M, First MB. *Instruction Manual for the Structured Clinical Interview for DSM-IV (SCID-IV). (SCID 1996 Revision)*. New York: Biometrics Research Department, New York Psychiatric Institute; 1996.
13. Schramm E, Hohagen F, Grasshoff U, et al. Test-retest reliability and validity of the Structured Interview for Sleep Disorders according to the DSM-III-R. *Am J Psychiatry.* 1993;150:867-72.
14. Edinger J, Kirby A, Lineberger M, Loiselle M, Wohlgemuth W, Means M. *The Duke Structured Interview for Sleep Disorders*. Durham, NC: Duke University Medical Center; 2004.
15. 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: Addison-Wesley; 1982:265-295.
16. Rechtshaffen A, Kales A. *A manual of standardized terminology, techniques, and scoring systems of sleep stages of human subjects*. Los Angeles: UCLA Brain Information Service/ Brain Research Institute; 1968.
17. Phillipson EA, Remmers JE. American Thoracic society Consensus conference on Indications and Standards for Cardiopulmonary Sleep Studies. *Am Rev Respir Dis* 1989;139:559-68.
18. American Sleep Disorders Association. EEG arousals: Scoring rules and examples - A preliminary report from the sleep disroders atlas task force of the American Sleep Disorders Association. *Sleep* 1992;15:173-84.
19. Mossman D, Somoza E. ROC curves, test accuracy and the description of diagnostic tests. *J Neuropsychiatry Clin Neurosci* 1991;3:330-3.
20. Zammit GK, McNabb LJ, Caron J, Amato DA, Roth T Efficacy and safety of eszopiclone across 6-weeks of treatment for primary insomnia. *Curr Med Res Opin* 2004;20:1979-91.
21. Fava M, McCall VW, Krystal AD, et al. Eszopiclone co-administered with fluoxetine in patients with insomnia co-existing with Major Depressive Disorder. *Biol Psychiatry* 2006;59:1052-60.
22. Roth T, Rogowski R, Hull S, et al. Efficacy and safety of doxepin 1 mg, 3 mg, and 6 mg in adults with primary insomnia. *Sleep* 2007;30:1555-61.
23. Scharf M, Rogowski R, Hull S, et al. Efficacy and safety of doxepin 1 mg, 3 mg, and 6 mg in elderly patients with primary insomnia: a randomized, double-blind, placebo-controlled crossover study. *J Clin Psychiatry* 2008;69:1557-64.
24. Lankford DA, Corser BC, Zheng YP, et al. Effect of gaboxadol on sleep in adult and elderly patients with primary insomnia: results from two randomized, placebo-controlled, 30-night polysomnography studies. *Sleep* 2008;31:1359-70.
25. Walsh JK, Salkeld L, Knowles LJ, Tasker T, Hunneyball IM Treatment of elderly primary insomnia patients with EVT 201 improves sleep initiation, sleep maintenance, and daytime sleepiness. *Sleep Med* 2010;11:23-30.
26. Krystal AD, Durrence HH, Scharf M, et al. Efficacy and safety of doxepin 1 mg and 3 mg in a 12-week sleep laboratory and outpatient trial of elderly subjects with chronic primary insomnia. *Sleep* 2010;33:1553-61.
27. Mayer G, Wang-Weigand S, Roth-Schechter B, Lehmann R, Staner C. Partinen M. Efficacy and safety of 6-month nightly ramelteon administration in adults with chronic primary insomnia. *Sleep* 2009;32:351-60.
28. Roth T, Soubrane C, Titeux L, Walsh JK; on behalf of the Zoladult Study Group. Efficacy and safety of zolpidem-MR: A double-blind, placebo-controlled study in adults with primary insomnia. *Sleep Med* 2006;7:397-406.
29. Carskadon MA, Dement WC, Mitler MM, Guilleminault C, Zarcone VP, Speigel R. Self-reports verses sleep laboratory findings in 122 drug-free subjects with complaints of chronic insomnia. *Am J Psychiatry* 1976;133:1382-8.
30. Buysse DJ, Reynolds CF, Hauri PJ, et al. Diagnostic concordance for DSM-IV sleep disorders: a report from the APA/NIH DSM-IV field trial. *Am J Psychiatry* 1994;151:1351-60.
31. Edinger JD, Wyatt JK, Stepanski EJ, et al. Testing the reliability and validity of DSM-IV-TR and ICSD-2 insomnia diagnoses. Results of a multitrait-multimethod analysis. *Arch Gen Psychiatry* 2011;68:992-1002.

ACKNOWLEDGMENTS

SUBMISSION & CORRESPONDENCE INFORMATION

DISCLOSURE STATEMENT

SUPPLEMENTAL MATERIAL

Table S1—Comparing home and lab ROC AUCs, optimal cutoff values, and accuracy of architectural and arousal-related PSG-based sleep variables from lab and home monitoring

Variable	Home				Lab			
	AUC	Value	Sens.	Spec.	AUC	Value	Sens.	Spec.
Sleep Episodes	0.564	86.25	0.531	0.526	0.609	83.25	0.567	0.571
STAGE 1	0.518	16.32	0.545	0.536	0.453	15.6	0.449	0.443
STAGE 2	0.551	195.55	0.564	0.557	0.517	196.22	0.500	0.505
REM	0.582	77.92	0.554	0.557	0.527	78.43	0.531	0.536
SWS	0.465	75.6	0.436	0.433	0.518	73.92	0.500	0.505

REM, rapid eye movement sleep; SWS, slow wave sleep.

Table S2—Comparing home and lab ROC AUCs, optimal cutoff values, and accuracy of PSG-based sleep onset insomnia variables from lab and home monitoring

Variable	Sleep Onset Insomnia Versus Normal Sleepers							
	Home				Lab			
	AUC	Value	Sens.	Spec.	AUC	Value	Sens.	Spec.
TST	0.545	374.78	0.553	0.551	0.523	375.05	0.553	0.551
SE	0.566	85.10	0.579	0.571	0.587	87.79	0.566	0.571
SOL	0.671	18.22	0.612	0.618	0.665	16.44	0.653	0.645
MWASO	0.525	45.27	0.510	0.513	0.637	41.43	0.571	0.566
TWASO	0.459	3.73	0.469	0.500	0.577	2.78	0.510	0.513

TST, total sleep time; SE, sleep efficiency; SO, sleep onset latency; MWASO, middle WASO; TWASO, terminal wake time

Table S3—Comparing home and lab ROC AUCs, optimal cutoff values, and accuracy of PSG-based sleep maintenance insomnia variables from lab and home monitoring

Variable	Sleep Maintenance Insomnia Versus Normal Sleepers							
	Home				Lab			
	AUC	Value	Sens.	Spec.	AUC	Value	Sens.	Spec.
TST	0.550	374.78	0.553	0.540	0.538	377.00	0.526	0.524
SE	0.600	84.28	0.592	0.587	0.665	86.21	0.632	0.635
SOL	0.599	16.78	0.556	0.553	0.592	16.00	0.619	0.605
MWASO	0.599	50.53	0.587	0.592	0.664	44.73	0.603	0.605
TWASO	0.519	4.05	0.540	0.526	0.611	3.42	0.556	0.553

TST, total sleep time; SE, sleep efficiency; SOL, sleep onset latency; MWASO, middle WASO; TWASO, terminal wake time.