THEORETICAL REVIEW

Insomnia with objective short sleep duration: The most biologically severe phenotype of the disorder

Alexandros N. Vgontzas a,*, Julio Fernandez-Mendoza a, Duanping Liao b, Edward O. Bixler b

Abstract

The diagnosis of insomnia, according to the Diagnostic and statistical manual of mental disorders, fourth edition, text revision (DSM-IV TR), is based solely on the subjective complaints of difficulty initiating or maintaining sleep, early awakening, and interrupted or non-restorative sleep and associated daytime functioning complaints. The International classification of sleep disorders (ICSD-2) also defines insomnia based on subjective complaints but includes polysomnographic (PSG) findings as diagnostic criteria, although not required for the diagnosis, for virtually all types of insomnia. However, there are currently no objective measures that are useful in the diagnosis, differential diagnosis, severity assessment, or therapeutic response of insomnia. Such a gap increases the uncertainty in diagnostic priorities and treatment guidelines and, to some extent, reinforces the perception among practitioners that insomnia is a rather trivial disorder from a medical standpoint. The goal of this review is to describe and discuss recent findings that demonstrate that: 1) insomnia is associated with significant cardiometabolic morbidity and mortality and 2) objective measures of sleep are useful in predicting the biological severity of the disorder.

Subjective sleep measures in the evaluation and treatment of insomnia

Self-reported quantitative assessment of various dimensions of sleep has been used in research and clinical practice for many years,
Abbreviations

ACTH adrenocorticotropic hormone
AUC area under the curve
CARDIA coronary artery risk development in young adults
CHD coronary heart disease
CRH corticotropin-releasing hormone
DSM-IV-TR Diagnostic and statistical manual of mental disorders, fourth edition, text revision
EEG electroencephalography
HPA axis hypothalamic-pituitary-adrenal axis
ICSD-2 International classification of sleep disorders, second edition
MMPI-2 Minnesota multiphasic personality inventory-2
MSLT multiple sleep latency test
OR odds ratio
PSG polysomnography
REM rapid eye movement
SDB sleep disordered breathing
SWS slow wave sleep
TST total sleep time
TWT total wake time
VO₂ overall oxygen consumption
WASO wake after sleep onset

and sleep diaries, for one to two weeks, are recommended as essential measures in insomnia research. Current quantitative diagnostic criteria for insomnia that are considered the most defensible include sleep onset latency or wakefulness after sleep onset of more than 30 min. However, individuals with insomnia tend to overestimate time to fall asleep or time of being awake after sleep onset, and to underestimate total sleep time in comparison to PSG measures. Thus, currently the proposed subjective sleep measures for insomnia are associated with unsatisfactory diagnostic accuracy and have limited generalizability in clinical practice. Furthermore, the validity of quantitative diagnostic cutoffs has not been tested in terms of discriminating individuals with and without clinically significant insomnia and associated medical morbidity risk.

Objective sleep measures in the evaluation and treatment of insomnia

The sleep laboratory is essential for the evaluation of patients with sleep disordered breathing (SDB) and the diagnosis of narcolepsy, and in the differential diagnosis of idiopathic vs. psychogenic hypersomnolence. In addition, sleep laboratory measurements provide valuable objective information on the initial effectiveness, continued efficacy or tolerance, and potential withdrawal effects of a hypnotic drug.

With the disorder of insomnia, the usefulness of the sleep laboratory has been at best controversial. While some investigators in the 70s and 80s considered that sleep laboratory evaluation was necessary for the complete assessment and diagnosis of insomnia, others believed that the vast majority of insomniacs, could be best assessed, diagnosed and treated in an office setting. In fact, the ICSD included PSG measures as diagnostic criteria, although not required, for virtually all diagnoses. Also, some studies reported that sleep laboratory findings resulted in a substantial modification of the initial clinical diagnostic formulation in more than half of the patients. For example, it was suggested that sleep apnea and nocturnal myoclonus (i.e., periodic leg movements) are frequent causes of chronic insomnia. Also, other studies suggested that rapid eye movement (REM) sleep latency was useful in separating insomniacs with depression. However, two studies published in 1998 and 2001 that evaluated the activity of the stress system in patients with insomnia led us to re-evaluate the potential usefulness of sleep laboratory measures in the evaluation of insomnia. Specifically, the findings of these two studies, suggested to us that objective short sleep duration in insomniacs may be very useful in the assessment of the biological severity of the disorder including its association with physiological hyperarousal and cardiometabolic morbidity.

Insomnia with objective short sleep duration is associated with activation of both limbs of the stress system

Stress has been associated with the activation of the HPA axis and the sympatho-adrenal-medullary axis, whereas corticotropin-releasing hormone (CRH) and cortisol (products of the hypothalamus and adrenals, respectively), and catecholamines (products of the sympathetic system) are known to cause arousal and sleeplessness to humans and animals. On the other hand sleep and particularly deep sleep has an inhibitory effect on the stress system including its main two components, the HPA axis and the sympathetic system. Given that insomnia is a sleep disorder associated with both cognitive-emotional and physiological hyperarousal, it was only natural for researchers to explore the association of this disorder with the stress system.

Until several years ago, few studies had assessed cortisol levels in insomniacs, and their results were inconsistent. The majority of these studies reported no difference between “poor” sleepers and normal individuals in the levels of 24-h cortisol and 17-hydroxysteroid excretion. A preliminary study by our group found that 24-h urinary free cortisol levels, norepinephrine, and catecholamine metabolites were positively correlated with PSG indices of sleep disturbance, i.e., total wake time (TWT) or wake time after sleep onset (WASO). Moreover, in a subsequent controlled study, 24-h serial adrenocorticotropic hormone (ACTH) and cortisol levels were significantly higher in insomniacs compared to normal sleepers (see Fig. 1). Within the 24-h period the greatest elevations were observed in the evening and during the first half of the night.
Furthermore, within the group of insomniacs, the sub-group with a high degree of objective sleep disturbance (% total sleep time < 70% of time in bed) had a higher amount of cortisol compared to the subgroup with a low degree of sleep disturbance (see Fig. 2). The cortisol levels for the group of insomniacs with high sleep efficiency were not different from normal controls. Based on these two studies, we concluded that in chronic insomnia: 1) the activity of the stress system is directly proportional to the degree of objective sleep disturbance, and 2) polysomnographic measures can provide a reliable index of the biological impact and severity of chronic insomnia.16,17

These findings were confirmed by several studies24–27 but not all.28,29 It appears that the difference between these two groups of studies is the degree of PSG documented sleep disturbance. For example, in the study by Rodenbeck et al.,24 the correlation between the area under the curve (AUC) of cortisol and sleep efficiency was –0.91, suggesting that high cortisol levels are present in those insomniacs with an objective short sleep duration. In contrast, in the study by Riemann et al.28 in which no cortisol differences were observed between insomniacs and controls, the objective sleep of insomniacs was very similar to that of controls (sleep efficiency of 88.2% vs. 88.6%). Furthermore, in a study that applied constant routine conditions,29 all indices of physiological arousal were increased but not to a significant degree due to lack of power and controls not being selected carefully.30 Interestingly, in this study a visual inspection of cortisol data suggested an elevation of cortisol values of 15–20% in the insomnia group, a difference which is similar to that reported in our study17 and is considered of clinical significance. Moreover, the study by Shaver et al.25 reported that objectively defined insomniacs had the highest levels of urinary cortisol as compared to both good sleepers and subjectively defined insomniacs.

Consistently with the studies of insomnia and HPA axis, other studies have shown higher sympathetic and central activation, including 24-h whole body and brain metabolic rate, heart rate, impaired heart rate variability, and increased catecholamine secretion, primarily in insomniacs who also had objective PSG sleep disturbance (see reviews by Bonnet and Arand31 and Riemann et al.32). For example, two early studies found small differences between subjectively defined good and poor sleepers in terms of increased heart rate,20,33 whereas two more recent studies have found significant changes in nocturnal heart rate and heart rate variability when comparing normal sleepers to insomnia patients who also met objective criteria of sleep disturbance.34,35 In a group of insomnia patients with PSG verified sleep disturbance, a significantly increased pupil size, indicative of sympathetic system activation, was observed compared to controls,36 whereas two other studies, in which insomnia diagnosis was based only on subjective measures, showed opposite results, i.e., decreased pupil size.37,38 Two studies by Bonnet and Arand39,40 investigated whole-body metabolic rate in insomnia patients with and without PSG documented sleep disturbance by measuring overall oxygen consumption (VO₂). VO₂ in insomnia patients with PSG documented sleep disturbance was constantly elevated at all measurement points as compared to carefully matched controls, whereas VO₂ was significantly increased in sleep state misperception insomnia patients compared to controls but to a lesser degree compared to insomnia patients with PSG documented sleep disturbance.39,40 Urinary circulating catecholamines and their metabolite levels have been found to be increased in insomniacs with objective sleep
disturbances or to correlate with PSG indices in insomniacs diagnosed solely on subjective complaints.\(^1\)\(^,\)\(^16\)\(^,\)\(^23\)\(^,\)\(^27\) Finally, two neuroimaging studies have shown that objective WASO is positively correlated with relative metabolism in the pontine tegmentum and in thalamocortical networks in a frontal, anterior temporal, and anterior cingulate distribution\(^41\) and negatively correlated with GABA levels in insomnia patients.\(^42\) Collectively, these studies yield the conclusion that physiological hyperarousal, i.e., hyperactivity of the HPA and sympatho-adrenal-medullary axes, is primarily present in insomnia patients with objective short sleep duration.

**Insomnia with objective short sleep duration is associated with medical morbidity and mortality**

Many studies have established that insomnia is highly comorbid with psychiatric disorders and is a risk factor for the development of depression, anxiety, and suicide.\(^1\) However, in contrast to the other most common sleep disorder, SDB, chronic insomnia has not been linked firmly with significant medical morbidity, e.g., cardiovascular disorders. Several questionnaire-based studies have shown a significant relationship between difficulty falling asleep or poor sleep and cardiac outcomes (see review by Bonnet and Arand\(^43\)). For example, in a prospective study from Japan,\(^44\) persistent (>4 years) complaints of difficulty initiating or maintaining sleep were associated with an increased risk of hypertension (odds ratio, OR = 1.96).\(^1\) In another prospective study, difficulty falling asleep (hazard ratio = 1.4) and difficulty staying asleep (hazard ratio = 1.3)\(^1\) were associated with acute myocardial infarction.\(^45\) Furthermore, previous studies have shown that sleep disturbances or complaints are associated with increased incidence of type 2 diabetes.\(^46\)–\(^48\) However, these studies did not include a PSG evaluation and could not control for sleep apnea or other sleep pathology; thus, their findings were dismissed as artifacts or flawed by many clinicians and researchers along.\(^50\),\(^51\) In fact, Kripke et al.\(^52\) found a reduced mortality rate for those individuals complaining of sleep difficulties after 6 y of follow-up.

Given the well-established association of hypercortisolemia with significant medical morbidity, i.e., hypertension, diabetes, metabolic syndrome, osteoporosis, and others, and our findings that insomnia with objective short sleep duration was associated with hypercortisolemia, we hypothesized that this type of insomnia will be associated with significant cardiometabolic morbidity and mortality.\(^16,\)\(^17\) Thus, we have examined in a systematic way the joint effects of insomnia and short sleep duration on the risk of hypertension, diabetes, neurocognitive impairment, and mortality in a general population sample randomly selected from Central Pennsylvania (the Penn State cohort).\(^53\),\(^54\)

The Penn State cohort is a population-based study of sleep disorders, which used a two-phase protocol in order to recruit participants from various age groups.\(^53\),\(^54\) In the first phase of the study, a sample of adult men and women (age ≥20 y) was randomly selected from local telephone households in Central Pennsylvania using the Mitofsky-Waksberg two-stage random digit dialing procedure. A within-household selection procedure described by Kish was used to select the specific man or woman to be interviewed. Telephone interviews were conducted with 4364 age-eligible men and 12,219 age-eligible women residing in the sample households, for a total sample of 16,583 with response rates of 73.5% and 74.1%, respectively. In the second phase of this study, a subsample of 741 men and 1000 women selected randomly from those subjects previously interviewed by telephone were studied in our sleep laboratory. The response rates for this phase were 67.8% and 65.8% for men and women, respectively. The design of the study included oversampling of those at higher risk for SDB and women with markedly higher levels of BMI to increase the precision of the risk estimates. Because of this sampling strategy, numeric sampling weights were developed for the analysis so that the estimates could be inferred to the general population. A comprehensive presentation of this sampling strategy has been presented elsewhere,\(^53\)–\(^55\) including the use of the NHANES III laboratory data as the standard to adjust both men and women in terms of sociodemographics to be representative of the national population.

Each subject selected for laboratory evaluation completed a comprehensive sleep history and physical examination. All subjects were evaluated for one night for 8 h (fixed-time period) in the sleep laboratory using PSG. Bedtimes were adjusted to conform to subjects’ habitual bedtimes, and subjects were recorded between 22:00 h – 23:00 h and 6:00 – 7:00 h. In this random general sample of Central Pennsylvania the vast majority of individuals went to bed between 22:00 h and 23:00 h, whereas only a small minority went to sleep outside of this time window and none for more than an hour. Thus, the maximum adjustment we had to do was of 1 h. The sleep recordings were subsequently scored independently, according to Rechtschaffen and Kales criteria. In the studies presented below sleep disordered breathing (SDB) was defined as an apnea or hypopnea index ≥5 events/h.

As part of this protocol we also assessed for the presence of all sleep disorders, based on a standardized questionnaire completed by the subjects on the evening of their sleep laboratory visit. The presence of sleep difficulty was established on three levels of severity. First, *insomnia* was defined by a positive response to the question “Do you feel you have insomnia?” with a duration of ≥1 y. Second, *poor sleep* was defined as a moderate to severe complaint of difficulty falling asleep (“Do you have difficulty falling asleep?”), difficulty staying asleep (“Do you have difficulty staying asleep?”), early final awakening (“Do you wake up in the morning earlier than desired?”), and/or non-restorative sleep (“Do you still feel groggy and unrefreshed after morning awakening?”). Third, *normal sleep* was defined as the absence of either of these two categories. Thus, normal sleep, poor sleep, and insomnia are three mutually exclusive groups as none in the poor sleep group reported having insomnia and none in the normal sleep group reported either insomnia or poor sleep.\(^55\)

**Insomnia with objective short sleep duration is associated with a high risk for hypertension**

We assessed the joint effect of insomnia and objective short sleep duration on hypertension risk in the Penn State cohort.\(^56\) In this study, PSG sleep duration was classified into three categories: \(≥ 6 \text{ h sleep (top 50% of the sample)}\); \(5 – 6 \text{ h (approximately the third quartile of the sample)}\); and \(< 5 \text{ h (approximately the bottom quartile of the sample)}\). Hypertension was defined based either on in-lab blood pressure measurements (i.e., diastolic blood pressure \(≥ 90 \text{ mm Hg and/or systolic blood pressure} ≥ 140 \text{ mm Hg}) or antihypertensive medication use. We controlled for age, race, sex, body mass index, diabetes, smoking, alcohol use, depression, SDB, and sampling weight.\(^1\) We found that compared

\(^{1}\) In logistic regression, the odds ratio is the ratio between the probability that an event occurs among one group of individuals vs. another and the probability that the event does not occur. It represents the statistical association between the group variable and the event of interest.

\(^{1}\) In survival analysis, the hazard ratio is the ratio of disease incidence among one group of individuals vs. another while accounting for the time to event or censoring. It represents instantaneous risk between the group variable and the event of interest over the study time period.

\(^{1}\) Sampling weight refers to the adjustment of the data for sample characteristics to ensure that the data is representative of the target population.
to the normal sleeping and the \( \geq 6 \) h sleep duration group, the highest odds of hypertension was in insomnia with \( \leq 5 \) h of sleep duration (OR = 5.1) and the second highest in insomniacs who slept 5–6 h (OR = 3.5). The odds for hypertension were significantly higher, but of lesser magnitude, in poor sleepers with short sleep duration. An objective sleep duration of 5–6 h was associated with a non-significant, slight increased risk of hypertension (OR = 1.1) whereas a sleep duration of \( \leq 5 \) h increased the risk by about fifty percent (OR = 1.56) compared with a group that slept more than 6 h. Importantly, the group of normal sleepers with short sleep duration did not show significant increased odds of hypertension (see Fig. 3). We concluded that insomnia with short sleep duration is associated with a high risk for hypertension, to a degree comparable to that of other common sleep disorders, e.g., SDB. At the same time, Lanfranchi et al.\(^57\) reported that nighttime systolic blood pressure was higher and day-to-night systolic blood pressure dipping was lower in insomniacs compared to controls and that the magnitude of beta EEG activity correlated with systolic blood pressure dipping in insomniacs. More recently, using longitudinal data from the Penn State cohort we have shown that insomniacs with short sleep duration are at a significantly higher risk of incident hypertension,\(^59\) suggesting that it is insomnia that causes hypertension and not vice versa. Cumulatively, these data indicate that objective short sleep duration may predict the biological severity of chronic insomnia, a prevalent condition whose medical impact has been apparently underestimated.

**Fig. 3.** Multivariable adjusted odds ratio of hypertension associated with insomnia with objective short sleep duration. Data are presented as odds ratio (OR) in bold, and 95% confidence interval (CI), in parentheses. Based on findings by Vgontzas et al.\(^56\)

**Insomnia with objective short sleep duration is associated with deficits in neuropsychological performance**

Although insomnia is associated with complaints of fatigue, difficulty in concentration, memory problems, and the like, objective neurocognitive impairment seems to be of subclinical nature in a variety of cognitive domains (see review by Fortier-Brochu et al.).\(^60\) We examined the joint effect of insomnia and objective sleep duration on neuropsychological performance, in the Penn State cohort.\(^61\) Individuals were classified as insomniacs or normal sleepers and were split according to PSG sleep duration into two categories: \(\geq 6\) h of sleep (“normal sleep duration”) and \(<6\) h of sleep (“short sleep duration”). We compared the groups’ performance on a comprehensive neuropsychological battery that measured processing speed, attention, visual memory and verbal fluency, while controlling for age, race, gender, education, body mass index, SDB, and physical and mental health problems.

No significant differences were detected between insomniacs and controls on either PSG variables or neurocognitive scores. However, there were significant interactions between insomnia and short sleep duration on specific neurocognitive tests. Specifically, the insomnia with short sleep duration group compared to the control with normal or short sleep duration groups showed poorer neuropsychological performance in variables such as processing speed, set-switching attention, and number of short-term visual memory errors and omissions. In contrast, the insomnia with normal sleep duration group showed no significant deficits when compared to controls. Based on these findings, we concluded that insomnia with objective short sleep duration is associated with deficits in set-switching attention/cognitive flexibility, a key component of the “executive control of attention”
Brain functions such as cognitive flexibility, working memory, and interference control have been linked to the activation of the prefrontal and anterior cingulate cortices. Tasks of set-switching attention also require the ability to retain both sets so they are ready to be recalled. In fact, neuroimaging studies show that set-switching attention and working memory cooperate in the same areas of the prefrontal cortex. These findings further support our hypothesis that objective sleep duration predicts the biological severity of chronic insomnia, including its effect on neurocognitive function.

The findings of our study are consistent with previous studies showing that performance in higher-order cognitive functions is impaired in insomniacs with objective sleep disturbances or correlates with objective markers of sleep severity, whereas performance in these cognitive functions is not impaired in insomniacs based solely on subjective criteria. In our study, the presence of a non-insomniac, short sleep duration group allowed us to demonstrate that neuropsychological deficits were associated to underlying physiological hyperarousal, a characteristic of chronic insomnia, rather than to short sleep per se. Combined, these studies suggest that insomnia with objective short sleep duration, via hypercortisolemia, affects brain structures and associated neurocognitive functions.

**Insomnia with objective short sleep duration and mortality**

Because our previous findings showed that insomnia with objective short sleep duration is associated with increased cardiometabolic morbidity, we further examined the effects of this insomnia subtype on all-cause mortality in the Penn State cohort. Participants were followed-up for 14 y (men) and 10 y (women). Individuals were classified as insomniacs or normal sleepers and were split according to PSG sleep duration into two categories: the “normal sleep duration group” subjects who slept ≥6 h and the “short sleep duration group” subjects who slept <6 h. We, again, adjusted for age, race, education, body mass index, smoking, alcohol, depression, SDB, and sampling weight.

The mortality rate was 21% for men and 5% for women. In men, mortality risk was significantly increased in insomniacs who slept less than 6 h compared to the “normal sleep duration” group (OR = 4.00) independent of comorbid conditions associated with mortality (see Fig. 6). Furthermore, there was a marginally significant trend toward higher mortality from insomnia and short sleep duration in patients with diabetes or hypertension (OR = 7.17) than in those without these comorbid conditions (OR = 1.45), suggesting that the impact of insomnia with short sleep duration was much stronger in those with diabetes and hypertension at baseline versus those who were healthy. In women, mortality was not associated with insomnia and short sleep duration which might be related to the fact that women were followed-up for a shorter period and that a smaller number were deceased at the time of follow-up compared to men. Interestingly, in the same cohort we failed to find any association between objective sleep duration and mortality regardless of insomnia status.

**Insomnia with objective short sleep duration: a disorder of sleep loss or of physiological hyperarousal?**

The findings of the physiological studies and of the epidemiological studies that suggest that insomnia with objective short sleep duration is associated with activation of both limbs of the stress system and significant cardiometabolic morbidity and mortality raise the question whether this type of insomnia is a disorder of sleep loss or of physiological hyperarousal. Data from several physiological domains support the latter.

For example, insomniacs complain that they are fatigued and sleepy during the day and one would expect that during the multiple sleep latency test (MSLT) they would demonstrate reduced sleep latencies. However, insomniacs compared to controls have either similar or increased daytime sleep latencies. Importantly, several studies have shown that, within insomniacs, those with shorter objective sleep duration show longer sleep latencies in the MSLT and are more alert in vigilance tests. This is in contrast to normal individuals who after modest short-term sleep loss experience significantly reduced sleep latencies on the MSLT and decreased alertness in vigilance tests, i.e., physiological sleepiness. Another example is the association of sleep deprivation and insomnia with peripheral cortisol levels. Normal individuals after total or partial sleep loss do not demonstrate increased levels of cortisol or sympathetic activity. Early studies that showed...
evening elevation of cortisol following several nights of sleep restriction were confounded by a stressful experimental paradigm.\textsuperscript{92,93} In contrast, insomniacs with objective short sleep duration demonstrate increased cortisol levels and sympathetic activity. These data combined suggest that the physiologic changes observed in insomniacs with short sleep duration are not a result of chronically accumulated sleep loss, but rather manifestations of an overall physiologic activation. In this regard, objective short sleep duration appears to be a marker of the physiologic hyperarousal in insomnia and not a measure of the cumulative chronic sleep loss.

**Sleep duration and cardiometabolic morbidity and mortality: quality versus length**

The previously reported studies on the synergistic effect of insomnia and short sleep duration make necessary to briefly examine the question whether sleep loss alone is a significant risk factor for cardiometabolic morbidity and mortality.

Sleep loss and sleep restriction are considered to be widespread in the general population and to have significant adverse health effects. Approximately one in three adult Americans are sleeping less than 7 h per night an amount at which physiological and neurobehavioral deficits manifest and become worse under chronic conditions.\textsuperscript{94} Indeed, the percentage of men and women reporting sleeping less than 6 h per night has increased significantly over the last 20 y. In animals, prolonged sleep deprivation is associated with death. In humans, sleep restriction is associated with major health risks from the cardiovascular and metabolic domains. Several experimental studies in the last decade have demonstrated that acute or short term sleep restriction is associated with imbalance of the glucose regulation, increased appetite, dysregulation of the appetite related hormones and increased inflammation.\textsuperscript{86,95,96} During the same period, many epidemiological studies have shown a consistent association between self-reported sleep duration and obesity and cardiometabolic morbidity and mortality.\textsuperscript{97} Interestingly, studies based on objective sleep duration, have reported inconsistent findings in terms of the association of sleep duration and cardiometabolic effects.\textsuperscript{95,98,99} This discrepancy may be explained by the fact that self-reported short sleep duration is influenced by sleep complaints, emotional and social stress, and unhealthy behaviors.\textsuperscript{100,101} This raises the possibility that self-reported short sleep duration does not reflect true sleep loss but rather is a marker of sleep complaints, such as insomnia and poor sleep, and chronic psychosocial stress.\textsuperscript{100,102,103} This possibility is further supported by our findings on insomnia and short sleep duration and cardiometabolic morbidity and mortality. For example, in our studies objective short sleep duration was associated with a significant risk for hypertension but this effect was much stronger and marked when we examined the joint impact of insomnia plus short sleep duration on blood pressure.\textsuperscript{56} Also in our mortality study, we found a strong association between insomnia with short sleep duration and mortality in men but we failed to find any association between objective sleep duration and mortality.\textsuperscript{78} Combined, these data suggest that in general population samples possibly the quality of sleep, i.e., being stressed and not satisfied with one's sleep is more important than the actual length of sleep. However, this does not negate the possibility that actual sleep loss, particularly in its severe form, may affect cardiometabolic morbidity and mortality.\textsuperscript{104,105} For example, studies using ancillary data to the coronary artery risk development in young adults (CARDIA) cohort have found that objective short sleep duration, as measured by actigraphy, is associated with increased risk of incident coronary artery calcification,\textsuperscript{104} incident hypertension as well as adverse changes in systolic and diastolic blood pressure over 5 y of follow-up.\textsuperscript{105} Also, an overnight PSG study in obese children aged 3–19 y found that children who slept less than 6 h during the PSG had higher fasting insulin and higher insulin resistance on the basis of the homeostatic model assessment method.\textsuperscript{106} Another recent study has found that short sleep duration, as measured by PSG, is associated with central obesity (i.e., waist circumference and abdominal diameter), a risk factor for cardiometabolic disorders, in a population sample of middle-aged women.\textsuperscript{107} Further studies, including both subjective and objective measures of sleep, should address this issue.

**Role of objective short sleep duration in the natural history of insomnia and poor sleep**

The prevalence of insomnia varies widely among epidemiologic studies ranging from 8 to 40%. A possible explanation for this wide variability is the different criteria used to define insomnia. For example, approximately 8–10% of the general population suffers from chronic insomnia,\textsuperscript{25,108,109} whereas another 20–30% of the population has insomnia symptoms, i.e., difficulties initiating sleep, difficulties maintaining sleep, early morning awakening, and/or non-restorative sleep, at any given time.\textsuperscript{108,109} In our recent studies, the term poor sleep refers to the presence in an individual of 1 or more insomnia symptoms without any duration criterion, whereas chronic insomnia is defined as a complaint of insomnia lasting more than a year (e.g., Bixler et al.\textsuperscript{55} and Vgontzas et al.\textsuperscript{56}). However, little is known about the longitudinal course of these two conditions and their bidirectional association, including shifting from one condition to another.

Insomnia is perceived to be a highly persistent disorder. However, risk factors associated with its persistence are not well understood. We have recently hypothesized that, if objective short sleep duration is a marker of the biological severity of the disorder, then it should predict also its persistence over time. Thus, we examined the role of PSG variables such as sleep duration or sleep apnea and other risk factors on the persistence of insomnia.\textsuperscript{110} From the 1741 individuals of the Penn State cohort, 1395 were followed-up after 7.5 y via telephone interview. PSG sleep duration was analyzed as a continuous variable and as a categorical variable: ≥6 h of sleep (“longer sleep duration”) and <6 h sleep (“short sleep duration”). The rates of insomnia persistence, partial remission, and full remission were 44%, 30%, and 26%, respectively. Objective short sleep duration significantly increased the odds of persistent insomnia as compared to normal sleep (OR = 3.19) and to fully remitted insomnia (OR = 4.92). Mental health problems at baseline were strongly associated with persistent insomnia as compared to normal sleep (OR = 9.67) and to a lesser degree compared to fully remitted insomnia (OR = 3.68). Sleep apnea did not predict the persistence of insomnia.

In two other studies we examined the clinical and PSG predic- tors of the incidence of chronic insomnia and the natural history of poor sleep in the general population.\textsuperscript{111,112} In the Penn State cohort, the incidence of poor sleep was 18%, the persistence and remission rates of poor sleep were 39% and 44%, respectively, and about one out of five (17%) subjects with poor sleep became chronic insomniacs. We found that both physical health conditions (e.g., obesity, sleep apnea, and ulcer) and mental health problems and malad- aptive personality traits as well as behavioral factors predicted the incidence of poor sleep, whereas the persistence of poor sleep was predicted by physical health conditions in combination with psychological distress. Importantly, we found that shorter objective sleep duration and a family history of sleep problems, but not sleep apnea, were risk factors for poor sleep evolving into chronic insomnia. These results suggested that objective short sleep dura- tion is a biologic marker of genetic predisposition to chronic insomnia.\textsuperscript{112}
Objective sleep duration, sleep misperception, and psychological profiles

According to the DSM-IV, individuals with chronic insomnia complain of sleep difficulties and frequently underestimate their sleep duration. Some investigators in the field of insomnia consider the underestimation of sleep duration a trait feature of all insomniacs, in which extreme cases might exist, whereas others suggest that a more severe subgroup of chronic insomnia patients who consistently underestimate their sleep duration deserves a separate diagnostic category. This latter view is represented by the ICD-2, which allows the diagnosis of "paradoxical insomnia" (so-called "sleep state misperception" or "subjective insomnia"), whereas the DSM-IV-TR does not include this diagnosis under "primary insomnia" because there is insufficient evidence to support its separate existence. To address this issue, we examined in the Penn State cohort whether objective sleep duration and psychological profiles, i.e., Minnesota multiphasic personality inventory-2 (MMPI-2), influenced the association between insomnia and sleep misperception. Individuals were classified as insomniacs and normal sleepers and were split into two groups based on their objective sleep duration: "normal sleep duration" (>6 h) and "short sleep duration" (<6 h).

Our findings showed that the degree of discrepancy between subjective and objective sleep duration was determined by two independent factors. Short sleepers reported more sleep than they objectively had, and insomniacs reported less sleep than controls with similar objective sleep duration. The additive effect of these two factors resulted in a significantly higher degree of underestimation of sleep time only in insomniacs with normal sleep duration which account for about 50% of all insomniacs in the general population. In fact, we found a significant interaction between insomnia and objective sleep duration on sleep misperception when the latter was defined as a clinically significant underestimation of sleep time (i.e., ≥−1.0 h). Interestingly, insomniacs with normal sleep duration showed an MMPI-2 profile of high depression and anxiety and low ego strength, whereas insomniacs with short sleep duration showed a profile of a medical disorder, i.e., depressive mood, fatigue, concerns about health, somatically focused anxiety, and poor health status (see Fig. 7). We concluded that sleep misperception, i.e., the underestimation of sleep duration, is prevalent among insomniacs with objective normal sleep duration, whereas insomniacs with objective short sleep duration are more likely to be accurate or to overestimate their sleep duration. Furthermore, anxious-ruminative personality traits and poor coping resources seem to mediate the underestimation of sleep duration.

The findings of our study are consistent with previous reports showing that sleep misperception or subjective insomnia is associated with psychological characteristics such as anxiety, rumination, worry, selective attention/monitoring toward sleep-related threats (see review by Harvey and Tang) but not with significant activation of both limbs of the stress system. Insomnia is associated with cortical arousal, as reflected by heightened higher frequencies (e.g., beta) in the EEG, higher rate of EEG arousals, and an inability to inhibit information processing. This heightened cortical arousal during sleep appears to be higher in insomniacs with normal sleep duration and may explain why insomniacs perceive their sleep as wake and as non-restorative. The specific pattern of cortical arousal in insomniacs with normal or short sleep duration needs to be further delineated in future studies.

In summary, cognitive-emotional and cortical arousal is present to a variable degree in all insomniacs. However, it is

![Fig. 7. Sleep misperception in insomnia: Role of objective sleep duration and psychological profiles. Panel A: data are presented as mean, in bold, and standard error (SE), in parentheses. Panel B: data are presented as mean T scores in each clinical and content scale (for the sake of simplicity SE have been omitted); 1-BS = hypochondriasis, 2- D = depression, 3-HY = hysteria, 4-PD = psychopathic deviate, 6-PA = paranoia, 7-PT = psychosis, 8-SC = schizophrenia, 9-MA = hypomania, 0-SI, social introversion, DEP = depression content scale, ANX = anxiety content scale, and ES = ego strength. Insomnia < 6 h (†) scored significantly higher in 1-BS, 2-BS, 3-HY, and DEP, whereas insomnia ≥ 6 h (‡) scored significantly higher in 1-BS, 2-BS, 4-PD, 6-PA, 7-PT, 0-SI, DEP, ANX, and lower in ES when compared to normal sleepers. Based on findings by Fernandez-Mendoza et al.]

Further evidence that insomnia with short sleep duration is the most severe phenotype of the disorder: heart rate variability, type 2 diabetes, and depression

Following our studies, there have been several reports with findings consistent with our model that insomnia plus objective short sleep duration has a significant effect on physical and mental health. Specifically, Spiegelhalder et al. published a study in which they determined the association of primary insomnia with heart rate and heart rate variability. Fifty-eight patients with primary insomnia and 46 healthy controls participated in the study. The study failed to show a difference in resting heart rate between the two groups and also could not confirm previous findings of an increase in sympathovagal balance and a decrease in parasympathetic nocturnal activity. However, when the authors restricted their analysis to insomnia patients with
objectively determined short sleep duration, they found reduced parasympathetic activity, as indicated by decreased high frequency power of heart rate variability (HRV), as well as decreased root mean square of successive RR index differences (RMSSD) and percentage of successive RRIs that differ by more than 50 ms (pNN50). Moreover, a recent study by De Zambotti et al.,126 has further supported that constant sympathetic hyper-activation, as measured by lower cardiac pre-ejection period (which is inversely related to sympathetic β-adrenergic activity), is present in insomniacs with polysomnographic sleep distur- bances. Consistent with the findings from these studies using PSG, a recent, large epidemiological study found a synergistic effect between subjective sleep disturbance and self-reported short sleep duration on the risk of coronary heart disease (CHD), so that the effect of short sleep duration on increasing CHD risk was greatest among those who reported sleep disturbances.127

In another study, Knutson et al.128 examined whether objective sleep duration and sleep quality were associated with markers of glucose metabolism among subjects with and without diabetes. This was an ancillary study to the CARDIA study. Sleep duration was estimated from 6 d of wrist actigraphy. Insomnia was defined as self-reported difficulty falling asleep or waking up in the night plus average sleep efficiency of <80% based on actigraphy. The study did not find any association between sleep measures and markers of glucose metabolism among subjects without diabetes. However, there was a significant association between sleep fragmentation and indices of glucose metabolism among patients with diabetes. Furthermore, a much stronger association was reported between objectively defined insomnia and markers of glucose metabolism. The authors concluded that patients with type 2 diabetes are at a higher risk of diabetes complications if the patients suffer from poor sleep quality, i.e., insomnia combined with objective short sleep duration.

A third study, by Troxel et al.129 examined whether insomnia, objectively measured sleep disturbances, and their interaction predict treatment outcome in depressed patients treated with psychotherapy or psychotherapy—pharmacotherapy combina-

tions. Participants were 711 depressed patients that have participated in clinical trials at the University of Pittsburgh. Objectively measured short sleep duration (total sleep time < 6 h) and prolonged sleep latency (>30 min) were derived from PSG sleep studies. Prolonged sleep latency alone or in combination with insomnia, predicted increased risk of non-remission. In addition, insomnia and short sleep duration, indi-

tually and in combination, were each associated with a significantly increased risk of non-remission. The authors concluded that objectively measured prolonged sleep latency and short sleep duration independently or in conjunction with insomnia are risk factors for poor depression treatment outcome and that insomnia combined with objectively measured sleep disturbance may represent a biologically more severe phenotype of insomnia.

**Habitual sleep duration and night-to-night variability**

In our studies, PSG-measured sleep duration is a strong predictor of HPA axis hyperactivity or medical morbidity among insomniacs.6,16,17,56,59,61,76 One of the frequent criticisms of our epidemiological studies is that objective sleep duration was based on one night of PSG, which may not be representative of the subjects’ habitual sleep duration. It should be noted that these studies investigated the relative sleep duration measured object-

ively (i.e., <6 h of objective sleep is relatively shorter than > 6 h of sleep) and that objective sleep duration was used as an internally valid marker of the severity of insomnia and not as the recommended optimum sleep duration for the general population. However, in previous studies, the association between objective short sleep duration and physiological hyperarousal (e.g., hyper-activity of the HPA axis, increased daytime MSLT sleep latency, and sympathetic activation) was based on 3 or 4 consecutive night sleep laboratory protocols, which should represent better the typical sleep profile of the subjects.16,17,19,31,132 Nevertheless, in-lab recordings of sleep tend to disturb a subject’s sleep (“first night effect”) and due to night-to-night variability of a subject’s sleep, multiple night recordings are required to obtain a repre-

tative sample of an individual’s typical habitual sleep pattern.6 It has been proposed that in-home recordings of 5 nights or more may provide a more accurate measure of a person’s typical sleep patterns,150 including an estimate of the night-to-night variability of sleep which appears to be greater in insomniacs.121 More recently, night-to-night variability assessed with actigraphy for 1—2 wk has been indicated as a useful marker of the severity of insomnia; i.e., there is a significantly larger night to night vari-

ability among chronic insomniacs than among controls.122 There was also a significant adverse association between intra-individual variability in sleep duration and fragmentation and psychosocial and physiological indices of stress, reflected as increased nighttime secretion of catecholamines,132,133 These studies suggest the potential usefulness of actigraphy to assess sleep patterns for a period of days or weeks in the “habitual home environment”.6,131,134 which is difficult with traditional PSG. However, several problems associated with the use of actigraphy, such as lack of an industry standard for the sleep algorithms used in different actigraphic devices and the propensity to over- or under-

determine sleep time, make its current use limited. Future studies using cost-effective, convenient methods should examine whether, for example, night-to-night variability in sleep duration may be a stronger predictor of cardiometabolic morbidity compared to average sleep duration.

**Other polysomnographic variables as potential markers of the biological severity of insomnia**

In our studies we have primarily focused on objective sleep duration as a marker of the biological severity of insomnia because cortisol levels, the main output of the stress system, showed their strongest association with short sleep duration.6,17 We have not systematically looked into other variables such as sleep latency or WASO. In our preliminary study on chronic insomnia and activity of the stress system, we reported that, for example, there was a significant correlation between WASO with 24-h urinary norepinephrine, dihydroxyphenylglycol, and dihy-

droxypheinylacetic acid.16 Also, we found significant correlations be-

tween percent stage 1 of sleep with the same molecules.16,17 Interestingly, Edinger et al.68 showed that PSG-measured WASO was the best single predictor of performance in tasks of set-switching attention in a large sample of chronic insomniacs. These results suggest that other variables of sleep efficiency and continuity may also serve as markers of the biological severity of the disorder. In fact, in one of our recent studies we demon-

strated that WASO was a significant predictor of who among poor sleepers will develop chronic insomnia.112 However, a potential disadvantage of biomarkers such as percent of stage 1 of sleep is that they require a full PSG study, whereas sleep duration could be obtained with simpler methods, e.g., actigraphy.

A sleep variable that deserves a separate examination is slow wave sleep (SWS). SWS is considered the best marker of sleep homeostatic mechanisms and its reduction indicates
compromise of sleep homeostasis.\textsuperscript{135} SWS at baseline is associated with lower activity of the sympathetic system and the HPA axis and after a night of total sleep deprivation or several nights of partial sleep restriction, increased SWS is associated with reduction of catecholaminergic activity and cortisol secretion.\textsuperscript{89,90,91} These findings have been extended into the clinical arena and recent studies have shown that experimental reduction of SWS results in insulin resistance and impaired glucose metabolism. Furthermore, more recent studies have shown that decreased SWS in elderly men increases the risk of developing hypertension.\textsuperscript{136}

In insomnia, the literature regarding SWS is inconsistent. While some studies have shown that SWS is decreased in insomniacs (see review by Pigeon and Perlis\textsuperscript{137}), the majority of the studies including those with a large number of participants have shown no differences between insomniacs and normal sleepers.\textsuperscript{11,69,79,121} These findings may indicate that there is possibly a group of insomniacs in which SWS is decreased. Some studies have explored the association between SWS and neurocognitive performance in insomniacs but results have been inconsistent.\textsuperscript{64,70} To date, no study has demonstrated that SWS is a predictor of cardiometabolic risks in insomnia. Further studies should focus on those insomniacs with lower SWS and examine whether this group is associated with higher morbidity and mortality.

**Phenotyping insomnia: diagnosis and treatment implications**

The field of sleep disorders medicine has attempted to define subgroups within insomnia based on etiology (i.e., primary versus secondary), age of onset (i.e., childhood versus adulthood), and discrepancy in objective versus subjective sleep findings.\textsuperscript{4} However, these subtypes show poor diagnostic reliability and have not been proven to be useful in terms of their impact on health or specific treatment outcomes.\textsuperscript{13,138,139}

The studies presented in this review have led us to suggest two phenotypes of chronic insomnia that are different in terms of etiology, pathophysiology, biological severity, psychological characteristics, diagnostic evaluation, and possibly, treatment response.\textsuperscript{40,47,56,59,61,78,110,112,115} The first phenotype is associated with cognitive-emotional, cortical, and physiological hyperarousal (i.e., short sleep duration and activation of the stress system), significant medical sequelae (e.g., hypertension, diabetes, neurocognitive deficits, increased mortality), and non-remitting course. The second phenotype is associated with cognitive-emotional and cortical arousal, but not with physiological hyperarousal (i.e., normal sleep duration and normal activity of the stress system), lack of significant medical sequelae, and is more likely to remit over time. Furthermore, the first phenotype is associated with a psychological profile typical of medical outpatients, whereas the second phenotype is associated with sleep misperception and an anxious-ruminative profile (see Fig. 8).

We do not know whether chronic insomnia with short sleep duration is in a continuum with insomnia with normal sleep duration or whether these two phenotypes are two distinct clinical entities. Based on our findings, we have recently proposed that poor sleep and chronic insomnia with normal sleep duration appear to be in a continuum with strong psychological roots, whereas poor sleep and chronic insomnia with objective short sleep duration appear to be in a continuum with a strong biological background (e.g., genetics and physiological hyperarousal).\textsuperscript{112} Furthermore, a recent study has provided for the first time evidence of the neural circuitry involved in stress-induced

---

![Image](119x107 to 467x368)

**Fig. 8.** Heuristic model of the underlying pathophysiological mechanisms and clinical characteristics of the two insomnia phenotypes based on objective sleep duration. The common characteristics of the two phenotypes are presented in the overlapping area, while their unique characteristics are presented in the areas of each phenotype that do not overlap.
insomnia associated with objective sleep disturbance in rats. However, insomnia with objective normal sleep duration (i.e., cognitive-emotional and cortical arousal and sleep misperception in the absence of signs of physiological hyperarousal) represents a challenge for animal models. Why some insomniacs with cognitive-emotional and cortical arousal do not exhibit physiological hyperarousal at the hypothalamic level, i.e., those with normal sleep duration, is not known. Previous studies have shown that age-related changes in sleep physiology increase the vulnerability of sleep to stress hormones (i.e., CRH administration). This biological vulnerability may play a significant role in the marked increase in the prevalence of insomnia in middle-aged and older adults. Future studies, including genetic, neurobiological, and longitudinal studies may shed light on the mechanisms of the biological vulnerability exhibited by insomniacs with short sleep duration.

Our findings on these proposed phenotypes may have a significant impact on how we diagnose and treat insomnia. In terms of diagnosis, the introduction of objective measures of sleep in the evaluation of insomnia may be of relevance for the clinician in terms of prioritizing intervention based on severity. Furthermore, these two phenotypes may respond differentially to treatment approaches. The first phenotype may respond better to treatments that primarily aim at decreasing physiological hyperarousal (e.g., cortisol) and increasing sleep duration, such as medication or other biological treatments, whereas the second phenotype may respond better to treatments that primarily aim at decreasing cognitive-emotional hyperarousal (e.g., rumination) and altering sleep misperception, such as sleep restriction, cognitive restructuring, behavioral experiments, or emotion regulation techniques. Supportive of the latter hypothesis is that cognitive-behavioral therapy for insomnia greatly improves subjectively but not objectively the sleep duration of insomniacs. Of course, it is possible that treatments that combine biological and psychological interventions may be the most effective for the more biologically severe phenotype, as it has been shown in other disorders (e.g., major depression). The differential treatment response of these two phenotypes should be tested in future placebo-controlled clinical trials.

Practice points

- In insomnia, objective sleep disturbance is directly proportional to the degree of the activity of both limbs of the stress system and other indices of physiological hyperarousal.
- Insomnia with objective short sleep duration is associated with significant cardiometabolic morbidity and mortality.
- Insomnia with normal sleep duration is associated with cognitive-emotional and cortical arousal and sleep misperception but not with signs of physiological hyperarousal or cardiometabolic morbidity.
- Objective measures of sleep duration may become part of the routine diagnosis and treatment of chronic insomnia in an office setting.
- These two phenotypes may respond differentially to treatment, i.e., insomnia with short sleep duration may respond better to biological treatments, whereas insomnia with normal sleep duration may respond better to psychological treatment alone.

Research agenda

In the future we need to:

- Study further the association of insomnia with short sleep duration with cardiometabolic changes, stress system, and immune system activity in experimental and longitudinal studies.
- Explore further the underlying genetic, neurobiological, and neuropsychological mechanisms of these two insomnia phenotypes.
- Explore the association of insomnia with short sleep duration with neurocognitive functioning and neurodegenerative disorders, e.g., mild cognitive impairment, amnestic and non-amnestic, dementia, or Parkinson’s disease.
- Examine the association of other sleep variables, such as sleep latency, wake after sleep onset, slow wave sleep, or nigh-to-night variability with the more biologically severe phenotype of insomnia.
- Replace the expensive, inconvenient, and impractical in-lab sleep measures with easy-to-use, inexpensive, home-based objective sleep measures.
- Study the efficacy of biological and psychological treatments, and their combination, on these two phenotypes of chronic insomnia.

Acknowledgments

This research was in part funded by the National Institutes of Health grants R01 51931, R01 40916 and R01 64415. The work was performed at the Sleep Research and Treatment Center at the Penn State University Milton Hershey Hospital, and the staff is especially commended for their efforts.

References


* The most important references are denoted by an asterisk.


