The Pathophysiology of Insomnia

Jessica C. Levenson, PhD; Daniel B. Kay, PhD; and Daniel J. Buysse, MD

Insomnia disorder is characterized by chronic dissatisfaction with sleep quantity or quality, that is associated with difficulty falling asleep, frequent nighttime awakenings with difficulty returning to sleep, and/or awakening earlier in the morning than desired. Although progress has been made in our understanding of the nature, etiology, and pathophysiology of insomnia, there is still no universally accepted model. Greater understanding of the pathophysiology of insomnia may provide important information regarding how, and under what conditions, the disorder develops and is maintained as well as potential targets for prevention and treatment. The aims of this report are (1) to summarize current knowledge on the pathophysiology of insomnia and (2) to present a model of the pathophysiology of insomnia that considers evidence from various domains of research. Working within several models of insomnia, evidence for the pathophysiology of the disorder is presented across levels of analysis, from genetic to molecular and cellular mechanisms, neural circuitry, physiologic mechanisms, sleep behavior, and self-report. We discuss the role of hyperarousal as an overarching theme that guides our conceptualization of insomnia. Finally, we propose a model of the pathophysiology of insomnia that integrates the various types of evidence presented.

ABBREVIATIONS: GABA = γ-aminobutyric acid; MnPO = median preoptic area; NREM = non-rapid eye movement; PEP = pre-ejection period; PSG = polysomnography; REM = rapid eye movement; SNP = single-nucleotide polymorphism; TMN = tuberomammillary nucleus of the posterior hypothalamus; VLPO = ventrolateral preoptic area

Insomnia disorder is characterized by dissatisfaction with sleep quantity or quality, associated with difficulty falling asleep, frequent nighttime awakenings with difficulty returning to sleep, and/or awakening earlier in the morning than desired.¹ Prevalence estimates of insomnia vary, with 30% to 43% of individuals reporting at least one nighttime insomnia symptom.³⁻⁶ Most reports suggest prevalence rates of insomnia disorder at 5% to 15%.⁴⁻⁶⁻⁷ Insomnia is a chronic problem in 31% to 75% of patients,¹⁻⁶⁻⁷ with more than two-thirds of patients reporting symptoms for at least 1 year.⁹ Although progress has been made in recent years regarding our understanding of the nature, etiology, and pathophysiology of

© 2015 AMERICAN COLLEGE OF CHEST PHYSICIANS. Reproduction of this article is prohibited without written permission from the American College of Chest Physicians. See online for more details.

DOI: 10.1378/chest.14-1617
insomnia, there is still no universally accepted model. This may be related to the heterogeneity of insomnia, its highly comorbid nature, or differences in what level of analysis the models use, from phenomenology to physiology. To be comprehensive, an etiologic or pathophysiologic model of insomnia should explain features such as the heterogeneity of symptoms and the risk insomnia confers for other comorbid conditions, such as depression and cardiometabolic syndrome. It should also explain the discrepancy between subjective (self-report) and objective (polysomnography [PSG]) measures of insomnia symptoms reported by some individuals with insomnia (see Reference 13 for a review). Greater understanding of the pathophysiology of insomnia may provide important information regarding how, and under what conditions, the disorder develops as well as potential targets for prevention and treatment. The aims of this review are (1) to summarize current breadth of knowledge on the pathophysiology of insomnia and (2) to present a model of the pathophysiology of insomnia that draws on evidence from various domains. Our article is intended to provide a brief overview of these topics for clinicians and researchers whose main focus is not insomnia. More extensive reviews of this topic can be found elsewhere.  

Levels of Analysis: An Approach to Understanding Insomnia

Although evidence-based assessments and treatments for mental disorders have been developed, diagnostic criteria for these conditions—including insomnia—are grounded in clinical consensus. Further progress depends on better understanding the etiology and pathophysiology of mental health problems. One framework for doing this has been offered by the National Institute of Mental Health's "Research Domain Criteria" initiative. While recognizing the value of current diagnostic categories, the National Institute of Mental Health has begun to emphasize observable "domains" of brain function pertinent to mental health. These research domains, such as positive emotion, negative emotion, and arousal, often show similar patterns of dysregulation across traditional diagnostic categories and can be examined across levels of analysis from genes to symptoms. These points pertain to insomnia as well. The first two editions of the International Classification of Sleep Disorders introduced > 25 diagnoses with "insomnia" as a cardinal symptom, but evidence for the reliability, validity, and distinct pathophysiology of these insomnia phenotypes has proved elusive. Partially as a consequence of this, both the International Classification of Sleep Disorders, Third Edition and Diagnostic and Statistical Manual of Mental Health Disorders, Fifth Edition, now propose a single major category for Insomnia Disorder or Chronic Insomnia Disorder. Nevertheless, there remains an impetus for the field to develop an evidence-based model of insomnia that accounts for heterogeneity in cause, symptoms, course, comorbidities, and consequences. This review considers evidence across seven levels of analysis based on the Research Domain Criteria framework: genetic, molecular, cellular, neuroanatomic, physiologic, behavioral, and self-report.

Hyperarousal: An Overarching Theme

Insomnia is often considered to be a disorder of hyperarousal, or increased somatic, cognitive, and cortical activation. Individuals with insomnia may experience physiologic hyperarousal in both central (cortical) and peripheral (autonomic) nervous systems (see References 20, 22, 23 for full review). Hyperarousal in insomnia can also refer to cognitive and emotional processes, with several theories suggesting that cognitive and affective hyperarousal at bedtime may contribute to both acute and chronic insomnia. Despite the frequent attention to hyperarousal in the literature, it is not frequently defined. In this report we conceptualize hyperarousal as heightened physiologic, affective, or cognitive activity, which interferes with the natural "disengagement from [...] the environment" and decreases the likelihood of sleep. Hyperarousal may be detected using such measures as increased cortisol, heart rate variability, EEG, or even self-report (eg, "I can't turn off my mind; "I feel so keyed up"). One of the challenges in identifying hyperarousal is that an individual always has some level of arousal, and the exact threshold for categorizing hyperarousal is not well defined. Thus, most studies have identified hyperarousal by noting differences between insomnia and control groups, rather than denoting a specific threshold. We propose hyperarousal as an overarching theme that, along with other contributory factors, enriches our understanding of the pathophysiology of insomnia at each level of analysis and across levels in an integrated model.

Genetics of Sleep and Insomnia

Sleep-wake traits, such as sleep duration and timing, are heritable and regulated by numerous genes. Animal and human studies also implicate genetic mechanisms in the etiology of insomnia. Seugnet et al isolated...
insomnia-like Drosophila flies (ins-1 flies) with rest-activity traits that resemble human insomnia, including decreased rest time, increased latency to a resting state after lights out, greater fragmentation of rest periods, and heightened activity levels. Whole-genome transcript profiling of ins-1 flies identified 755 genes with human homologs that were differentially expressed compared with wild-type flies. Genes found to be conserved in ins-1 flies are associated with sensory perception, metabolism, cell surface signaling, and neuronal activity and may have implications for understanding the genetics of human insomnia.

Most genetic studies in humans have used a limited set of self-report items to categorize insomnia symptom phenotypes and yielded a wide range in heritability estimates for insomnia (h² = 0-81%) across family history and twin studies. Studies using more stringent criteria to define insomnia have produced more realistic and reliably modest h² estimates ranging from 31% to 58%. Candidate gene studies have identified gene variants that may be involved in the pathophysiology of insomnia, including ApoE, PER3, HLA DQB1*0602, homozygous Clock gene 3111C/C Clock, and short (s-) allele of the 5-HTTLPR. A genomewide association study found numerous single-nucleotide polymorphisms (SNPs) significantly associated with insomnia symptoms. The most significant SNPs occurred within genes involved in neuroplasticity (eg, ROR1, PLCB1, EPHA4, and CACNA1A), stress reactivity (eg, STK39, USP25, and MARP10), neuronal excitability (eg, GABRB1 and DLG2), and mental health (eg, NPAS3).

Overall, current evidence suggests significant heritability and multigene involvement in the pathophysiology of insomnia. Genes linked to brain functioning, arousal regulation, and sleep-wake processes have been most consistently found to be associated with insomnia. The complex interplay of these genes may account, at least in part, for the heterogeneity observed in insomnia symptoms and consequences. Future genetic studies with detailed assessment of sleep and health history of patients with chronic insomnia disorder may further refine our understanding of genetic factors involved in the development and characteristics of insomnia.

Molecular Mechanisms of Sleep and Insomnia

Numerous sleep regulatory substances are linked to circadian rhythmicity and sleep regulation. Although recognizing the oversimplification, we argue that endogenous molecules can be categorized as primarily wake-promoting/sleep-suppressing (eg, catecholamines, orexin, and histamine) and sleep-promoting/wake-suppressing substances (eg, γ-aminobutyric acid [GABA], adenosine, serotonin, melatonin, prostaglandin D2). Very few molecular studies have been conducted in insomnia and have focused on only a limited set of molecules (eg, cortisol and GABA). Table 1 lists studies linking various molecules to insomnia. Findings are mixed across studies, and no consistent pattern for a specific type of molecule (sleep vs wake promoting) has emerged. Despite contradictory evidence, results have largely been interpreted within the context of the hyperarousal hypothesis. For example, increased and decreased GABA in the occipital cortex of patients with insomnia have been reported to be consistent with the hyperarousal model of insomnia. However, sleep regulatory molecules interact with each other in complex ways (described in more detail later), and many of their effects are dependent on the milieu of the brain state; that is, they are state-dependent. These factors make it highly unlikely that all cases of insomnia can be explained by alterations in any single type of molecule (eg, hyperarousal-related). A more sophisticated conceptualization holds that chronic insomnia results from disintegration of the alternating rhythms of wake-promoting and sleep-regulatory molecules in the brain. Constant routine, in-home PSG, and sleep-deprivation studies, particularly those that examine wake- and sleep-promoting molecules (or their mRNA and associated micro-RNA) during different states across the 24-h day, may prove fruitful in elucidating the molecular underpinnings of chronic insomnia. In addition, studies that more fully link this level of analysis with the genetic underpinnings are also needed.

Cellular Mechanisms of Sleep and Insomnia

Many of the molecules involved in sleep-wake regulation are produced by specific brain structures with widespread projections throughout the brain. There is, however, mounting evidence that many sleep regulatory molecules affect neurons locally, in the regions in which they are produced. In local sleep theory proposed by Krueger et al, sleep is defined as a fundamental emergent property of highly interconnected neurons, or cortical columns. Local sleep propensity and slow wave amplitude are posited to be dependent on accumulation of sleep-regulatory substances (eg, tumor necrosis factor-α and IL-1β) resulting from prior neuronal use. Synchronous firing within cortical columns is postulated to propagate slow wave activity in adjacent regions through humoral and electric interactions,
leading eventually to a “global” sleep state in the entire organism.

From this perspective, insomnia may not be a “whole-brain” event (ie, a simple matter of imbalance between global amounts of sleep and wake). An animal model of insomnia has demonstrated simultaneous localized Fos activation in both sleep-promoting and wake-promoting regions during global sleep.60 In humans, spectral EEG methods have identified heightened regional electrical brain activity in patients with insomnia during non-rapid eye movement (NREM) sleep.61,62 Merica et al61 proposed that the lack of objective sleep disruption in many patients with insomnia may be due to isolated neuronal groups remaining active during PSG-defined sleep. This dynamic in the brain may be experienced as wakefulness by many patients with insomnia and miscategorized as “normal” sleep based on standard PSG criteria.63 Advances in neuroimaging technology would be needed to determine whether insomnia is associated with a distributed pattern of wakefulness at the neuronal level or is better characterized by a region-specific persistence of wake-like brain activity during globally defined EEG sleep more consistent with the next level of analysis.64

### TABLE 1

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Method</th>
<th>Insomnia vs Control Subjects</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcium</td>
<td>Blood serum levels</td>
<td>↑</td>
<td>41</td>
</tr>
<tr>
<td>γ-Aminobutyric acid</td>
<td>Average brain spectroscopy</td>
<td>↓</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Occipital cortex spectroscopy</td>
<td>↑</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Anterior cingulate and occipital cortex spectroscopy</td>
<td>↓</td>
<td>46</td>
</tr>
<tr>
<td>Melatonin</td>
<td>Evening wake/early sleep blood serum levels</td>
<td>↓</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Urinary excretion</td>
<td>Shifted</td>
<td>48</td>
</tr>
<tr>
<td>Noradrenaline</td>
<td>Urinary excretion</td>
<td>↓</td>
<td>49</td>
</tr>
<tr>
<td>Corticotropin-releasing hormone</td>
<td>Blood serum levels</td>
<td>↑</td>
<td>50</td>
</tr>
<tr>
<td>Adrenocorticotropic hormone</td>
<td>Blood serum levels</td>
<td>↓</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>24-h blood plasma levels</td>
<td>↑</td>
<td>51</td>
</tr>
<tr>
<td>Cortisol</td>
<td>Evening and morning salivary levels</td>
<td>↑ (Evening), ns (morning)</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Blood serum levels</td>
<td>ns</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Salivary levels</td>
<td>ns</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Evening and morning salivary levels</td>
<td>ns (Evening), ↓ (morning)</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Evening and morning salivary levels</td>
<td>ns</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Evening wake/early sleep blood plasma levels</td>
<td>↑</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Evening/early sleep blood plasma levels</td>
<td>↑</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Blood serum levels</td>
<td>↑</td>
<td>50</td>
</tr>
</tbody>
</table>

↓ = insomnia less than control subjects; ↑ = insomnia greater than control subjects; ns = no significant difference.

Sleep-Wake Regulation and Neural Circuitry of Sleep

On the global level, sleep is regulated by coordinated wake and sleep brain networks. Insomnia may plausibly involve dysregulation within these networks.

**Wake Systems and Hyperarousal in Insomnia**

The major wake-promoting systems of the brain include the “bottom-up” reticular activating system, limbic networks, and the “top-down” cognitive systems. Neural systems originating in the brainstem, thalamus, and hypothalamus65-67 constitute the ascending reticular activating system. This system projects to the cortex via the thalamus and basal forebrain and includes cholinergic pedunculopontine and laterodorsal tegmental nuclei, noradrenergic locus coeruleus nuclei, serotonergic dorsal and median raphe nuclei, the parabrachial nucleus, the histaminergic tuberomammillary nucleus of the posterior hypothalamus (TMN), and basal forebrain cholinergic nuclei. Orexin/hypocretin neurons of the lateral hypothalamus project to all of the arousal-promoting centers in the brainstem and hypothalamus and reinforce their activity. Emotional and cognitive systems can enhance
monoaminergic expression and lead to suppression of sleep-promoting regions such as the ventrolateral preoptic area (VLPO). Inputs to the arousal system may suppress the firing of VLPO neurons, disinhibiting the orexin/hypocretin and TMN neurons and thereby opposing sleep pressure.

During initiation of normal sleep, arousal systems are down-regulated by inhibition from the VLPO and median preoptic area (MnPO). The activation of arousal centers at the end of the sleep period is sufficient to terminate sleep. Activity of arousal systems (eg, cortisol) responsible for alertness is modulated by the circadian timing system. Insomnia is often considered a disorder of excessive activation of the arousal systems of the brain (ie, hyperarousal). Hyperarousal in the physiologic, emotional, or cognitive networks is believed to prevent sleep regulatory processes from naturally occurring in patients with insomnia (see References 20, 22-25). However, other evidence suggests that hyperarousal is neither necessary nor sufficient for the development of chronic insomnia. For example, many patients with insomnia show no signs of cardiovascular, body temperature, or cortisol marker of hyperarousal, and not all individuals with these common markers of hyperarousal develop insomnia.

Sleep Systems: Two-Process Model of Sleep Regulation

The activity of arousal and sleep centers described above is modulated by two critical physiologic processes: wake-dependent (homeostatic) sleep drive and circadian rhythmicity. These two processes have been described in the two-process model and related conceptualizations, such as the opponent-process model of sleep-wake regulation. According to the two-process model of sleep regulation, sleep propensity is regulated by the interaction of a wake-dependent process (process S) and a relatively wake-independent circadian process (process C). Process S dictates that greater brain use during wakefulness increases sleep need and is measured by greater θ activity in the waking EEG and higher amplitude EEG power in the δ range (0.5-4.5 Hz) during NREM sleep. On the network level, sleep onset is driven by activation of GABAergic and galanin neurons in the VLPO and MnPO. Accumulation of extracellular adenosine during prior wakefulness has been posited as a primary input to these systems. Axons from VLPO/MnPO send outputs to arousal centers in the hypothalamus and brainstem, inhibiting arousal-promoting neurons of TMN, dorsal and median raphe nuclei, and locus coeruleus while simultaneously promoting sleep. Circadian sleep propensity is regulated by intrinsic circadian oscillations governed by the suprachiasmatic nuclei of the hypothalamus. Exogenous light, melatonin, and social factors can influence the suprachiasmatic nuclei-regulated circadian processes in the body, such as REM sleep, body temperature, and endogenous melatonin. Optimal sleep is believed to occur when the S- and C-processes driving sleep are appropriately coordinated.

One hypothesis based on the two-process model is that insomnia results from insufficient sleep propensity during the desired sleep period because of dysfunction in the S- or C-process. Evidence is mixed on whether patients with insomnia compared with control subjects have less robust slow wave activity following sleep deprivation suggestive of a deficiency in process S. Evidence linking insomnia to markers of circadian dysfunction, including delayed or advanced core body temperature rhythms, or increased mean nocturnal core body temperature in different insomnia phenotypes, suggest dysregulation of process C. However, some studies failed to find an association between core body temperature and insomnia.

Sleep Switch

The mutually inhibitory circuitry of the VLPO and arousal centers of the brain is often described as a central “flip-flop switch” regulating the activity of wake and sleep promoting systems to produce bistable sleep-wake states. From this perspective, sleep and wake states are achieved via reciprocal inhibition between the VLPO/MnPO regions and monoaminergic brainstem and hypothalamic arousal centers (see Reference 67 for review). Overriding the flip-flop switch has been proposed as a mechanism of insomnia. Although maintaining heightened awareness in the presence of sleep debt may be necessary and beneficial in rare times of crisis, insomnia may result from chronic coactivation of sleep and wake circuits during the desired sleep period. This conceptualization of insomnia is consistent with an animal study showing that the reciprocal inhibitory innervation between the VLPO and the arousal system can decouple under stressful conditions, resulting in a unique state with simultaneous sleep and wake features. According to this model, the core feature of insomnia is not reduced sleep or excessive wakefulness but rather the simultaneous activation of brain structures responsible for each state.
Other authors have proposed that insomnia is an unstable state in which individuals rapidly transition in and out of sleep-wake states (ie, a flickering switch). Individuals with greater subjective-objective sleep discrepancy may have more brief arousals from sleep (eg, Reference 79) and more frequent sleep-wake transitions during the sleep onset interval, consistent with rapid switching between sleep and wake states.

**Structural and Functional Neuroimaging**

Although slow wave activity during NREM sleep is commonly believed to be global and homogeneous in the cerebrum, numerous studies have demonstrated that sleep is a dynamic process in space and time. Compared with a resting wake state, NREM sleep is associated with lower whole-brain metabolism, particularly in the cortical association areas. More specifically, increased slow wave activity during NREM sleep corresponds with reduced cerebral blood flow globally but most strongly correlates with decreases in brain regions involved in adapting behavior to environmental pressures (ventrolateral prefrontal cortex), conscious processes (anterior cingulate, precuneus/upper cuneus, mediobasal preoptic area), action selection (basal ganglia), and generation of slow oscillations during slow wave sleep (brainstem, midbrain, and thalamic structures). Principal components analysis of PET scan data identified two brain networks associated with sleep: (1) reduced blood flow in frontal and parietal association cortices and hippocampus, and increased flow in the cerebellum; and (2) reduced blood flow in the thalamus and a region that, on visual inspection, overlaps with the precuneus and cuneus.

Lesion and structural neuroimaging studies also suggest that specific brain regions may be associated with insomnia. Animal studies demonstrated that lesions of the thalamus, raphe nucleus, or mediobasal preoptic area result in insomnia. von Economo observed the sleep of patients affected by the encephalitis pandemic of 1918, and these observations revealed an association between insomnia and lesions in the anterior hypothalamus. In addition, patients with traumatic brain injuries who endorse insomnia symptoms had overlapping lesions in the left dorsomedial frontal cortex. Structural imaging studies have identified reduced gray matter volume in left orbitofrontal, prefrontal, precuneus, and temporal cortices in patients with insomnia. These structures may represent dysfunctional nodes in networks of sleep/wake regulation.

Functional imaging studies suggest that patients with insomnia have smaller reductions in brain activity during NREM sleep relative to resting wake. Specifically, the frontoparietal cortex, medial temporal lobes, thalamus, anterior cingulate, precuneus, and brain stem arousal networks have been implicated. Corsi-Cabrera et al also examined the topographic distribution of brain wave activity associated with wakefulness as an index of cortical activation during the sleep onset period among individuals with primary insomnia. They found higher β activity in left frontal and frontal midline regions during W and N1 and higher levels of temporal coupling linking the frontal, parietal, and posterior midline regions during the sleep onset period in primary insomnia as compared with control subjects. One overarching hypothesis is that the regional patterns of greater activation during sleep in patients with insomnia reflects impaired deactivation and disengagement of brain regions involved in executive control, attention, and self-awareness and may contribute to the experience of insomnia.

**Electrophysiologic and Physiologic Dysregulation in Insomnia**

Hyperarousal has been examined using various electrophysiologic (EEG) and physiologic measures during sleep and wakefulness. EEG indicators of hyperarousal include increased high-frequency EEG activity (β and γ), decreased δ activity, and increased REM EEG arousals. As discussed later, physiologic measures include increased body temperature, skin resistance, metabolic rate, and heart rate, among others.

**NREM Sleep Instability**

The Neurocognitive Model of insomnia posits that acute insomnia may be perpetuated by maladaptive behavioral coping strategies and may develop into chronic insomnia as a result of conditioned arousal. Conditioned arousal is the repeated association of sleep-related cues with wakefulness and/or arousal, which, over time, results in an arousal response when a sleep-related stimulus is presented. The neurocognitive model focuses specifically on cortical arousal as the mechanism underlying chronic insomnia, as indexed by high-frequency EEG activity (β and γ, 16–50 Hz). This EEG activity is hypothesized to increase around sleep onset as a result of classic conditioning (ie, a learned response to cues associated with sleep). Data to support this hypothesis show diminished δ and increased high-frequency NREM EEG power among patients with insomnia, as well as an association between high-frequency EEG activity and subjective sleep complaints. High-frequency EEG
enhances sensory and information processing, which may contribute to the subjective-objective discrepancy that often characterizes insomnia.\textsuperscript{21} Evidence is mixed regarding insomnia-control differences in high-frequency EEG activity during wakefulness.\textsuperscript{72,100} However, high-frequency waking EEG activity significantly correlates with high-frequency EEG activity during NREM\textsuperscript{100} and with self-reported hyperarousal symptoms.\textsuperscript{72} These findings support the hypothesis that high-frequency EEG power in insomnia is a marker of CNS hyperarousal.\textsuperscript{101}

**REM Sleep Instability**

The REM sleep instability model\textsuperscript{102} hypothesizes that the subjective experience of insomnia is related to decreased REM sleep percent and increased REM EEG arousals.\textsuperscript{103} In one study, arousals and awakenings during REM more precisely distinguished patients with insomnia from good sleepers than NREM parameters.\textsuperscript{102} Fragmented REM sleep may promote the perception of increased wakefulness and nonrestorative sleep in insomnia, which may contribute to subjective-objective sleep discrepancies insomnia.\textsuperscript{102}

**Physiologic Hyperarousal**

Additional support for the involvement of hyperarousal in the pathophysiology of insomnia comes from findings examining various measures of physiologic arousal among individuals with insomnia. As early as 1967, Monroe\textsuperscript{104} showed that poor sleepers had increased body temperature, vasoconstrictions, body movements, and skin resistance as compared with good sleepers. Insomnia has also been associated, in some studies, with increased 24-h metabolic rate (as measured by oxygen consumption),\textsuperscript{27,105} 24-h adrenocorticotropic hormone and cortisol levels,\textsuperscript{71} and heart rate.\textsuperscript{106} Some investigators have demonstrated greater inhibition of facial muscle activity and increased cardiac vagal tone in response to sleep-related emotional stimuli among individuals with insomnia as compared with good sleepers.\textsuperscript{107} Others have specifically examined the sleep-onset period, finding increased frontalis electromyogram, increased heart rate, and decreased finger temperature among subjects with insomnia as compared with control subjects up to the point of sleep onset.\textsuperscript{108} Findings have also shown sympathetic activation among patients with insomnia during sleep onset, as evidenced by consistently lower cardiac pre-ejection period (PEP) values as compared with good sleepers.\textsuperscript{106,109} Cardiac PEP is the time interval from the beginning of ventricular depolarization (marked by the onset of the QRS complex in the ECG) to the opening of the aortic valve. PEP duration is inversely related to $\beta$-adrenergic tone.\textsuperscript{106} Thus, lower PEP values indicate enhanced activation of the sympathetic nervous system.

Many studies investigating physiologic arousal and hyperarousal in insomnia have included small samples and have not been consistently replicated.\textsuperscript{52} For this reason, it is not possible to specify diagnostic thresholds for any single physiologic measure in insomnia. Noting these findings, Varkevisser et al\textsuperscript{52} cautioned against overemphasizing hyperarousal in the conceptualization of chronic insomnia. Although hyperarousal is an important heuristic concept in many models of insomnia, several conflicting points remain to be resolved. First, the extent to which hyperarousal is a cause or a consequence of insomnia has not been elucidated. Second, targeting hyperarousal in the treatment of insomnia (eg, through relaxation training) is often less effective than approaches that enhance sleep processes (eg, sleep restriction and hypnotic medications). Third, markers of hyperarousal can also be interpreted as insufficient inhibition of arousal by sleep-promoting processes.

**Behavioral and Cognitive Contributions to Insomnia**

Various biologic mechanisms regulate sleep and contribute to insomnia. Behavioral and cognitive mechanisms (ie, beliefs that contribute to specific behaviors) can also regulate sleep and contribute to, and exacerbate, insomnia.

**Perpetuating Factors**

The “3P model”\textsuperscript{118} a diathesis-stress model, describes a set of predisposing, precipitating, and perpetuating factors that may contribute to the development and maintenance of insomnia. Predisposing factors, such as age or sex, make an individual more susceptible to insomnia, and precipitating factors are events that coincide with the onset of insomnia, such as major stressors.\textsuperscript{111} Perpetuating factors, the largest focus of the 3P model, are behaviors and beliefs that maintain insomnia,\textsuperscript{111} such as increasing time in bed to “catch up” on sleep.\textsuperscript{110} However, extended time in bed perpetuates insomnia because it leads to increased wakefulness, fragmented sleep, variability in sleep timing,\textsuperscript{110} and associations between the sleep environment and wakefulness. Thus, initial attempts to reduce symptoms of insomnia may evolve into perpetuating factors themselves.
Stimulus Control

In 1972, Bootzin\textsuperscript{112} proposed that stimuli associated with sleep (eg, a quiet, dark bedroom) become discriminative stimuli that reinforce sleep. Insomnia may result from inadequate sleep-promoting stimuli or from the presence of stimuli that are antithetical to sleeping;\textsuperscript{112} such as phone calls, reading, or worry. Stimulus control therapy for insomnia aims to separate the stimuli associated with sleep from the stimuli associated with other activities.\textsuperscript{112}

Cognitive Model

The cognitive model of insomnia\textsuperscript{25} proposes that individuals with insomnia are susceptible to excessive worry and unpleasant intrusive thoughts, particularly those related to getting enough sleep and the consequences of sleep disturbance. This worry may develop into sleep-related anxiety, lead to increased vigilance for sleep-related threats (eg, watching the clock at night), and ultimately result in an exaggeration of the magnitude of the actual sleep disruption. Cognitive therapy for insomnia challenges these maladaptive cognitive processes and limits the behaviors that maintain unhelpful beliefs and insomnia.\textsuperscript{113}

In the psychobiologic inhibition model,\textsuperscript{24,114} a variant of the cognitive model, sleep is thought of as automatic,\textsuperscript{24} whereas insomnia is thought of as a failure of automatic sleep. The model specifically focuses on the attention-intention-effort pathway as one sleep inhibitory process, in which three processes occur: (1) increased selective attention to sleep and symptoms of insomnia; (2) an increase in the subjective value of sleep, which may contribute to explicit “intention to sleep”; and (3) gradual development of increased effort to sleep, described as “sleep effort syndrome.”\textsuperscript{114} Thus, treatment of insomnia should focus on cognitive strategies that aim to reverse sleep-related attentional bias\textsuperscript{113} and behavioral strategies\textsuperscript{115} aimed at reducing sleep effort.

Self-Report and Insomnia

Unlike many other sleep disorders, insomnia disorder relies on self-report for diagnosis; physiologic markers of sleep dysregulation (PSG) or hyperarousal are not routinely indicated for the evaluation of insomnia.\textsuperscript{116–118} Several groups have developed self-report measures aimed at detecting insomnia and assessing insomnia-related experiences and impairments. Focus groups that captured patients’ subjective experience of insomnia highlighted the pervasive impact of the disorder, the perception that others do not fully understand the impact of insomnia, and the importance of daytime symptoms of insomnia.\textsuperscript{119} Self-report measures that were developed to assess presleep thought content\textsuperscript{120} and sleep-related quality of life impairment\textsuperscript{121} in insomnia demonstrated that (1) presleep cognitive activity among patients with insomnia focuses on rehearsal/planning, sleep and its consequences, and autonomic experiences, among others, and (2) the most common impairments pertain to energy/motivation, performance at work, cognitive functioning, and emotion regulation. This work has also shown that self-report symptom scales reliably discriminate individuals with insomnia and good sleepers.\textsuperscript{120,122}

Although PSG usually shows abnormalities in sleep architecture and continuity among individuals with insomnia, the magnitude of patient-control differences is often small, and the severity of objective findings is often less than that obtained by self-report.\textsuperscript{1} Nevertheless, differences among patients with insomnia in subjective and objective measures may have important clinical implications. For instance, a meaningful discrepancy between subjective and objective measurement of sleep is prevalent among patients with insomnia with objective normal sleep duration but not among patients with insomnia with objective short sleep.\textsuperscript{123} Moreover, compared with patients with insomnia who have normal overall sleep duration, patients with insomnia with short objectively measured sleep duration are at increased risk for adverse health outcomes, including hypertension, diabetes, physiologic hyperarousal, cognitive difficulties, and even mortality.\textsuperscript{124,125} The magnitude and night-to-night variability of discrepancy between self-reported and objectively measured sleep may itself constitute a high-risk presentation that may be informative for examining the etiology and pathophysiology of insomnia.\textsuperscript{126,127} Thus, objective assessments of sleep may be useful adjunctive measures in predicting the biologic severity and medical impact of insomnia, and cost-effective objective measures of sleep should be considered in the standard diagnostic procedure for insomnia to differentiate phenotypes.\textsuperscript{124,125}

Integration and Treatment Implications

Our intent in this article has been to outline the mechanisms by which insomnia develops and is maintained, highlighting findings in the literature at various levels of analysis. Table 2 summarizes the evidence at each unit of analysis, which indicates that evidence for the pathophysiology has been generated using numerous methodologies based on a range of theoretical scientific perspectives. Integration of the evidence presented here allows us to propose one possible model for the
### TABLE 2  | Evidence for the Pathophysiology of Insomnia at Each Unit of Analysis

<table>
<thead>
<tr>
<th>Unit of Analysis</th>
<th>Evidence for This Perspective in Pathophysiology of Insomnia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genes</strong></td>
<td>Elevated family risk for insomnia</td>
</tr>
<tr>
<td></td>
<td>Elevated genetic risk in twin studies</td>
</tr>
<tr>
<td></td>
<td>Insomnia phenotype in drosophila related to mutations in 755 genes with human homologs</td>
</tr>
<tr>
<td>Candidate gene studies support association between aspects of insomnia and ApoC-4, PER3/4, HLA DQB1*0602, 3111C/C Clock, short (s-) allele of the 5-HTTLPR</td>
<td></td>
</tr>
<tr>
<td>Numerous SNPs identified in human genomewide association study studies</td>
<td></td>
</tr>
<tr>
<td><strong>Molecules</strong></td>
<td>Mixed findings regarding the role of wake- and sleep-promoting molecules in insomnia; no consistent pattern for a specific type of molecule has emerged</td>
</tr>
<tr>
<td></td>
<td>Unlikely that all cases of insomnia can be explained by alterations in any single molecule type</td>
</tr>
<tr>
<td></td>
<td>See Table 1 for links between various molecules and insomnia</td>
</tr>
<tr>
<td><strong>Cells</strong></td>
<td>Simultaneous localized Fos activation in both sleep-promoting and wake-promoting regions during global sleep in rats. Individual cortical columns show sleep-like activity while other parts of the brain show wake-like activity</td>
</tr>
<tr>
<td></td>
<td>Neuronal use results in modulation of gene expression in sleep regulatory substance, which acts locally in the brain to promote sleep</td>
</tr>
<tr>
<td><strong>Circuits</strong></td>
<td>In animals, lesions of the anterior ventral and the dorsomedial thalamus, raphe nucleus, or paramedial preoptic area results in insomnia</td>
</tr>
<tr>
<td></td>
<td>Less robust slow wave activity following sleep deprivation among subjects with insomnia than control subjects (process S); delayed and advanced core body temperature rhythms and heightened nocturnal core body temperature linked to insomnia (process C)</td>
</tr>
<tr>
<td></td>
<td>Those reporting greater subjective-objective sleep discrepancy demonstrate sleep-related behaviors consistent with rapid switching between sleep and wake states</td>
</tr>
<tr>
<td></td>
<td>Reduced gray matter volume in left ventromedial prefrontal cortex, precuneus, and temporal cortices in patients with insomnia</td>
</tr>
<tr>
<td><strong>Physiology</strong></td>
<td>Patients with insomnia have smaller reductions in brain activity during NREM sleep relative to resting wake</td>
</tr>
<tr>
<td></td>
<td>Higher β activity in left frontal and frontal midline regions during W and N1, higher levels of temporal coupling linking the frontal, parietal, and posterior midline regions during the sleep onset period in insomnia as compared with control subjects</td>
</tr>
<tr>
<td></td>
<td>Diminished δ and increased high-frequency NREM EEG power among patients with insomnia</td>
</tr>
<tr>
<td></td>
<td>Association between high-frequency EEG activity and subjective sleep complaints</td>
</tr>
<tr>
<td></td>
<td>High-frequency waking EEG activity correlates with high-frequency EEG activity during NREM, and with self-reported hyperarousal symptoms</td>
</tr>
<tr>
<td></td>
<td>Arousals and awakenings during REM sleep more precisely distinguished subjects with insomnia from good sleepers than the NREM parameters</td>
</tr>
<tr>
<td></td>
<td>Insomnia associated with increased body temperature, vasoconstrictions, body movements, skin resistance, 24-h metabolic rate, 24-h ACTH, cortisol levels</td>
</tr>
<tr>
<td></td>
<td>Continuous sympathetic hyperactivation among subjects with insomnia during sleep onset</td>
</tr>
<tr>
<td><strong>Behavior</strong></td>
<td>Some efficacious treatments for insomnia focus on resolving the behavioral and cognitive factors contributing to and exacerbating insomnia. These include:</td>
</tr>
<tr>
<td></td>
<td>Increasing the association between the bed and being asleep</td>
</tr>
<tr>
<td></td>
<td>Reestablishing a consistent sleep-wake schedule</td>
</tr>
<tr>
<td></td>
<td>Restricting time in bed to increase sleep drive and, subsequently, sleep efficiency</td>
</tr>
<tr>
<td></td>
<td>Reducing somatic tension or intrusive thoughts that are antithetical to sleep</td>
</tr>
<tr>
<td></td>
<td>Psychotherapy targeting maladaptive beliefs about sleep</td>
</tr>
<tr>
<td><strong>Self-reports</strong></td>
<td>Self-report measures discriminate subjects with insomnia from good sleepers</td>
</tr>
<tr>
<td></td>
<td>Presleep mentation among subjects with insomnia focuses on rehearsal/planning, sleep and its consequences, and autonomic experiences</td>
</tr>
<tr>
<td></td>
<td>The areas of energy/motivation, performance at work, cognitive functioning, and emotion regulation are the most commonly reported sleep-related impairments in insomnia</td>
</tr>
</tbody>
</table>

*(Continued)*
pathophysiology of insomnia. Figure 1 depicts this model, in which insomnia is most likely to develop in those who have increased genetic risk and who experience abnormalities in neurobiological processes. These trait-like vulnerabilities may lead to neurophysiologic hyperarousal and to psychologic and behavioral processes, which, individually or together, increase an individual’s risk for developing insomnia and associated downstream health consequences. Precipitating stressors and other person-specific factors (e.g., age, sex) moderate these relationships. The extent to which an individual with insomnia shows evidence of abnormalities in each of the processes depicted may vary among different individuals.

Accordingly, interventions aimed at preventing or resolving symptoms of insomnia may target various aspects of the identified pathophysiologic processes. The current gold standard psychologic treatment of insomnia is Cognitive Behavioral Therapy for Insomnia,\textsuperscript{128–130} which is typically composed of multiple treatment elements: stimulus control therapy, sleep restriction therapy, relaxation training, cognitive therapy, and sleep hygiene education (see Reference 131 for a review). These treatment elements focus on: (1) increasing the association between the bed and being asleep; (2) reestablishing a consistent sleep-wake schedule; (3) restricting time in bed to increase sleep drive and, subsequently, sleep efficiency; (4) reducing somatic tension or intrusive thoughts that are antithetical to sleep; (5) targeting maladaptive beliefs about sleep; and (6) maintaining good sleep practices.\textsuperscript{131} As reviewed by Buysse\textsuperscript{132} and Morin and Benca,\textsuperscript{7} the efficacy of Cognitive Behavioral Therapy for Insomnia for the treatment of insomnia has been demonstrated. Although these approaches make efforts toward reducing cognitive and emotional arousal, there has also been a call for therapies that target

\textbf{TABLE 2 (continued) Evidence for This Perspective in Pathophysiology of Insomnia}

<table>
<thead>
<tr>
<th>Unit of Analysis</th>
<th>Evidence for This Perspective in Pathophysiology of Insomnia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underestimation of sleep duration is prevalent among those with a subjective complaint of insomnia but objectively normal sleep duration</td>
<td>The magnitude of discrepancy between self-reported and objectively measured sleep may itself constitute a high risk factor for insomnia.</td>
</tr>
</tbody>
</table>

\textit{ACTH} = adrenocorticotropic hormone; \textit{NREM} = non-rapid eye movement; \textit{SNP} = single-nucleotide polymorphism.

Figure 1 – Model of the pathophysiology of insomnia. GABA = \(\gamma\)-aminobutyric acid; SNP = single-nucleotide polymorphism.
physiologic hyperarousal across the night and day. Some studies have documented changes in physiologic measures with pharmacologic therapies, but much more work is needed to draw conclusions about the efficacy of these approaches.

In addition, several efficacious pharmacologic treatments for insomnia target various aspects of the identified pathophysiologic processes (see References 133, 134 for review). For example, benzodiazepine receptor agonists (eg, temazepam, zolpidem), which are generally effective in the treatment of insomnia, promote sleep by enhancing widespread inhibitory activity of GABA. The tricyclic drug doxepin has shown efficacy for sleep initiation and maintenance insomnia; although not US Food and Drug Administration-approved, trazodone is also widely used for insomnia. The sedative effect of these drugs is mainly achieved by targeting the histaminergic arousal system. More recently, there has been progress in the development of orexin receptor antagonists for the treatment of insomnia (eg, almorexant, suvorexant); suvorexant was recently approved by the Food and Drug Administration for this indication. These drugs target the orexin system that promotes arousal of brainstem/hypothalamic arousal centers. Other medications have been suggested for the treatment of insomnia, but additional work is needed to demonstrate safety and efficacy. Future work on the pathophysiology of insomnia may help to identify specific mechanisms in specific patient groups, which may lead to more targeted pharmacotherapy.

Conclusions

Evaluating evidence from a range of domains and across various levels demonstrates not only the advances made in understanding the pathophysiology of insomnia but also the areas in which additional support is needed and the type of analysis that might fill a gap in the literature. For example, human genetic studies that include more refined sleep measures (ie, clinical assessment, sleep diaries, actigraphy, PSG) may provide greater power needed to identify the role of genes in the pathophysiology of insomnia. At the neural and physiologic levels, future work should examine the potential role of impaired “switching processes” and local and circuit-level sleep dysregulation. Functional imaging, high-density EEG, magnetoencephalography, and magnetic resonance spectroscopy are useful tools for such studies. Last, additional studies at all levels of analysis should more deeply examine subjective-objective discrepancies as one clue to insomnia pathophysiology. Greater understanding of the pathophysiology of insomnia may provide important information regarding new targets for prevention and treatment.

Acknowledgments

Financial/nonfinancial disclosures: The authors have reported to CHEST the following conflicts of interest: Dr Levenson receives royalties from American Psychological Association books and receives grant support from the American Psychological Foundation. Dr Buysse has served as a consultant for Merck & Co, Inc; Medscape; Purdue Pharma LP; Emmi Solutions, LLC; Eisai Co, Ltd; CME Outfitters, LLC; and Otsuka Pharmaceutical Co, Ltd. Dr Kay has reported that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

Role of sponsors: The sponsor had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.

References

25. Harvey AG. A cognitive model of insomnia.


21. Perlis ML, Giles DE, Mendelson WB, Bootzin RR, Wyatt JK. 


42. Clinton JM, Davis CJ, Zielinski MR, Jewett KA, Krueger JM. Biochemical regulation of sleep and sleep biomarkers. 

Griffith LC. Neurmodulatory control of sleep in Drosophila melanogaster: integration of competing and complementary behaviors. 


Plante DT, Jensen JE, Schoering I, Winkelman JW. Reduced y-aminobutyric acid in occipital and anterior cingulate cortices in primary insomnia: a link to major depressive disorder? 


Xia L, Chen GH, Li ZH, Jiang S, Shen J. Alterations in hypothalamic-pituitary-adrenal/thyroid axes and gonadotropin-releasing hormone in the patients with primary insomnia: a clinical research. 


Varkevisser M, Van Dongen HP, Kerkhof GA. Physiologic indexes in chronic insomnia during a constant routine: evidence for general hyperarousal? 

Backhaus J, Junghanss K, Hohagen F. Sleep disturbances are correlated with decreased morning awakening salivary cortisol. 


Roddenbeck A, Huether G, Rüther E, Hajak G. Interactions between evening and nocturnal cortisol secretion and sleep parameters in patients with severe chronic primary insomnia. 


Krueger JM, Rector DM, Roy S, Van Dongen HP, Belenky G, Panksepp J. Sleep as a fundamental property of neuronal behaviors. 


114. Espie CA, Broomfield NM, MacMahon KM, Macphee LM, Taylor LM. The attention-intention-effort pathway in the development of...


