The Past and Future of Psychiatric Sleep Research

Thomas Pollmächer 1,2

1 Center of Mental Health, Klinikum Ingolstadt, D-85049 Ingolstadt, Germany; thomas.pollmaecher@klinikum-ingolstadt.de; Tel.: +49-841-880-2200
2 Department of Psychiatry and Psychotherapy, Ludwig Maximilians University, D-80336 Munich, Germany

Abstract: Sleep studies in psychiatric disorders date back to the first half of the 20th century. So far, success in establishing disease-specific sleep-related biomarkers has been quite limited. This is particularly obvious regarding insomnia, where there is no reliable correlation between subjective complaints and physiological measures of sleep. Finally, it must be acknowledged that the physiology of sleep and wakefulness and their subjective perception are essentially independent dimensions. Still, however, these dimensions are mixed up in clinical practice and research. This creates confusion, can be harmful for patients, and is identified as a major obstacle for successful psychiatric sleep research. It is proposed here that future sleep research should treat physiological sleep as a variable in psychiatric disorders which, independently of patient perception, has a transdiagnostic value, as was already proposed a decade ago by the Research Domain Criteria.

Keywords: sleep; psychiatric; wakefulness

1. Introduction

Since antiquity, disturbed sleep has been closely linked to mental disorders. Difficulties in initiating and maintaining sleep, as well as reduced sleep quality, are among the most prominent and most impairing complaints of psychiatric patients. Throughout history, from Aristotle via Hippocrates to Kraepelin [1], mood disorders, in particular, have been associated with sleep disturbances. These have collectively been termed “insomnia” probably since the 17th or 18th century [2] and since then, discussion has continued on whether insomnia is a mere consequence of certain conditions and disorders or a nosological entity in itself.

For centuries, scientific knowledge on sleep and its disorders relied exclusively on patients’ reports, and in few cases on visual observation. This changed fundamentally a few years after the discovery of the electroencephalogram (EEG) by the German psychiatrist Hans Berger in the late 1920s, which allowed for the continuous measurement of brain activity [3]. Soon, it became obvious that along with changes in EEG waves during sleep, numerous other physiological measurements (e.g., respiration, muscle tone, eye movements and others) showed systematic variation across the sleep–wake cycle. The discovery of rapid eye moments (REM) during sleep in the 1950s [4] enabled the conceptualisation of REM and non-REM sleep as distinct vigilance states within sleep, which are differentially regulated and have distinct functions [5]. Parallel studies led to the two-process model of sleep regulation, explaining the influence of circadian and homeostatic components on the duration, depth and timing of sleep [6].

Among clinicians, increasing knowledge on physiological sleep regulation particularly excited psychiatrists, because sleep studies promised to open a window into the brain and help gather objective information supplementary to subjective reports obtained from patients. The hope for objective indicators of psychiatric illness seemed to be fulfilled for the first time with the seminal findings of David Kupfer and the Pittsburgh group, published in The Lancet, who reported reduced latencies to rapid eye movement sleep in depressed patients [7].
In the following decades, psychiatric sleep research became very popular. Many aspects of sleep were investigated by polysomnography and hundreds of studies were published on numerous disorders (see for example the review by Baglioni et al., 2016 [8] and Table 1). However, PSG studies never entered clinical practice to support the diagnosis, treatment or prevention of psychiatric disorders. Even the most recent diagnostic systems (DSM-5, ICD-11 and ICSD-3) rely exclusively on patient reports when psychiatric sleep disturbances (including so-called insomnia disorder) are concerned.

Table 1. Changes in PSG parameters in psychiatric disorders. Data according to Baglioni et al. [8].

<table>
<thead>
<tr>
<th>Mental Disorder</th>
<th>Sleep Continuity</th>
<th>Sleep Stages</th>
<th>REM-Sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Efficiency</td>
<td>Latency</td>
<td>Nots</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major depression</td>
<td>↓</td>
<td>↓</td>
<td>↑</td>
</tr>
<tr>
<td>Post traumatic stress disorder</td>
<td>↓</td>
<td>n.s.</td>
<td>↑</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>↓</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Borderline personality disorder</td>
<td>↓</td>
<td>n.s.</td>
<td>↑</td>
</tr>
<tr>
<td>Anorexia nervosa</td>
<td>↓</td>
<td>n.s.</td>
<td>-</td>
</tr>
<tr>
<td>Attention deficit disorder</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

↓ = Decrease, ↑ = Increase, - = No data available for evaluation, n.s. = not significant.

The present article will first address the reasons for the failure to establish sensitive, objective sleep parameters of specific clinical value in the past. Then, I will argue that despite this historical failure, it is promising to further pursue psychiatric sleep research, although from a slightly different perspective.

2. Psychiatric Sleep Research until the Turn of the Millennium

Following the discovery of REM sleep, psychiatrists made increasing use of polysomnography as a window to the brain. Sleep studies were performed in most if not every disorder mentioned in the chapter “Mental and Behavioral Disorders” of the International Classification of Diseases (ICD). Dementias, alcohol and substance abuse disorders, schizophrenia and related disorders, affective disorders, anxiety disorders and personality disorders were investigated. The bottom lines of hundreds of papers published during these decades were as follows: (1) Disturbed sleep is extremely prevalent in patients with psychiatric disorders. (2) Sleep parameters which can be subjectively estimated and objectively measured, such as sleep onset latency, total sleep time, number of nocturnal awakenings, etc., show very high within- and between-subject variability. Correlations between corresponding subjective and objective measures are low. (3) Some objective aspects of NREM and REM sleep show prominent changes in some disorders; however, none of these changes are of sufficient sensitivity and specificity to be used as a reliable diagnostic marker. Ruth Benca and colleagues summarised the results of studies performed up to the early 1990s in an excellent review [9]. They mentioned that the most prominent alterations were found in affective disorders, mainly depression. Some of these relate to NREM, particularly slow-wave sleep, but those most widely studied are related to REM sleep measures. It seems worthwhile to elaborate somewhat more on these.

3. REM Sleep Pressure in Depression

The seminal finding of reduced REM sleep latency in depression, published in 1972 [7], opened two major avenues for future research: first, to establish specific clinical markers for diagnostic and predictive purposes and second, to obtain a valid target for experimental pathophysiological research. I will not further elaborate here on the second avenue. But, just to give one prominent example, in the framework of the aminergic–cholinergic
imbalance theory of depression, cholinergic substances have been used to challenge the REM regulation system in healthy people and depressed patients (e.g., [10]).

Quite soon after the description of reduced REM latency, it became obvious that this marker was not highly specific for depression (e.g., [9,11]), nor was this the case for increased REM density or increased amount of REM sleep [8]. Still, however, these three parameters indicating increased REM sleep pressure turned out to be quite informative.

For example, it has been shown that 75% of patients who displayed a shortened REM latency during a depressive episode relapsed within two years, whereas this was the case for only about 30% of those whose REM latency was normal [12]. Large PSG studies, in particular the “Munich vulnerability study”, were performed in families including multiple members affected with depression [13]. These studies were highly demanding because hundreds of depressed patients had to be screened to end up with reasonable numbers of healthy first-degree relatives of affected patients (high-risk probands, HRPs). Still, it was shown that these healthy HRPs displayed an increased REM density. Moreover, following up these individuals for several years revealed that they were at increased risk of developing an affective disorder [14], indicating that increased REM pressure might be a bio- or trait-marker of affective disorder [15]. However, the number of subjects included in these follow-up studies was small, weakening the empirical strength of the findings.

Unfortunately, these promising results on increased REM sleep pressure as a predictor of depressive illness were not further pursued, thus being neither confirmed nor disproven until now, although very recently, one big PSG study of 1820 community-dwelling individuals pointed once more to increased REM density as a trait marker of depression [16].

In general, however, studies using PSG in psychiatric patients have consistently declined following the turn of the millennium (e.g., for major depression, from 76 PubMed-listed papers between 1991 and 2000 to 47 from 2011 to 2020).

A number of reasons may explain this prominent decline in interest. (1) PSG studies are expensive, both in terms of technical needs and human resources. (2) PSG studies are demanding for the study subjects, which limits the willingness to participate, particularly among psychiatric patients. (3) PSG results are quite sensitive to psychotropic drugs and psychotherapy, which makes rigorous study designs indispensable. (4) Around the turn of the millennium, brain imaging and molecular genetics became increasingly popular and “trendy” in psychiatry, absorbing almost all public interest and funding.

All these reasons added to the disappointment with the limited clinical usefulness of PSG studies in psychiatric disorders. As the “story” of REM sleep in depression shows, this disappointment is only partially justified. Moreover, the presumed failure may be based on a quite fundamental misconception, as I will discuss below.

4. The Case of Insomnia Disorder

As already mentioned, since the term “insomnia” was coined some 300 years ago, there has been an ongoing debate on whether patients’ complaints about the quantity and quality of sleep are a mere consequence of different conditions or whether insomnia may represent a distinct nosological entity. The answer to this question might be approached empirically or in a normative way.

Normatively, insomnia has been defined as a distinct disorder in the ICD. In the 10th edition [17], it was listed among psychiatric disorders and called “non-organic” insomnia (ICD code F51.0), as opposed to the neurological disorder “organic insomnia” (ICD code G47.0). The Diagnostic and Statistical Manual of Psychiatric Disorders used quite a different term, “primary insomnia”, up to its 4th edition [18]. The DSM-5 [19] then used the term “insomnia disorder”, which was taken up by the 11th edition of the ICD [20]. Moreover, the ICD-11 removed insomnia from the context of mental and behavioural (i.e., psychiatric) disorders and moved it to a separate chapter for sleep disorders, chapter 07 (the International Classification of Sleep Disorders, 3rd edition (ICSD-3), also uses the term “insomnia disorder”).
From an empirical perspective, however, it is obvious that the symptoms of insomnia disorder, as listed in the DSM-5 (there are slight differences between the definitions in the DSM-5, ICD-11 and ICSD-3 regarding the symptoms of insomnia, but those are not relevant in the present context) (see Table 2, criteria A to E), can be found in a wide range of conditions from all medical specialities. Many of them (e.g., endocrine disorders, infection and inflammation, sleep apnoea syndrome, narcolepsy, medication effects, effects of alcohol and drugs) have proven, well-established negative effects on sleep and also a clear-cut negative impact on PSG measures. For others, particularly other psychiatric disorders, PSG effects are more debated, as discussed above.

Table 2. DSM-5 insomnia disorder criteria.

A. A predominant complaint of dissatisfaction with sleep quantity or quality, associated with one (or more) of the following symptoms:
1. Difficulty initiating sleep. (In children, this may manifest as difficulty initiating sleep without caregiver intervention.)
2. Difficulty maintaining sleep, characterised by frequent awakenings or problems returning to sleep after awakenings. (In children, this may manifest as difficulty returning to sleep without caregiver intervention.)
3. Early-morning awakening with inability to return to sleep.
D. The sleep difficulty is present for at least 3 months.
B. The sleep disturbance causes clinically significant distress or impairment in social, occupational, educational, academic, behavioral, or other important areas of functioning.
C. The sleep difficulty occurs at least 3 nights per week.
F. The insomnia is not better explained by and does not occur exclusively during the course of another sleep-wake disorder (e.g., narcolepsy, a breathing-related sleep disorder, a circadian rhythm sleep-wake disorder, a parasomnia).
H. Coexisting mental disorders and medical conditions do not adequately explain the predominant complaint of insomnia.
G. The insomnia is not attributable to the physiological effects of a substance (e.g., a drug of abuse, a medication).

Nosological systems (for example, see again the DSM-5, Table 2) separate this range of conditions affecting sleep from insomnia disorder in quite a fuzzy way. Insomnia disorder should not be diagnosed if:
- Another sleep–wake disorder better explains the insomnia or it occurs exclusively during the course of another sleep–wake disorder (criterion F);
- The insomnia is attributable to the physiological effects of a substance (criterion G);
- Coexisting mental disorders and medical conditions adequately explain the predominant complaint of insomnia (criterion H).

Hence, none of these conditions categorically excludes the diagnosis of insomnia disorder. Because there are no operationalised definitions of a “better” or “adequate explanation”, and as it remains unclear when insomnia is “attributable” to the physiological effects of a substance or not, it is left to each clinician or researcher whether he labels the insomnia complaints of a patient with, for example, major depression, alcohol dependence or sleep apnoea syndrome as symptoms of the respective condition or, alternatively, as a separate, than so-called “comorbid”, insomnia disorder. It is quite obvious that this nosological approach is problematic both for clinical work and for research on disturbed sleep.

However, before turning to the problems caused by the insomnia disorder concept, I would like to acknowledge some positive consequences. First of all, subjective complaints about sleep are extremely frequent and, for numerous patients, they are also extremely disturbing, particularly if they go along with impaired daytime functioning. The concept of insomnia disorder takes those patients seriously, more so than the view that problems in initiating and maintaining sleep are just subordinate symptoms of another condition does. The concept of insomnia disorder also emphasises subjectively disturbed sleep as an important therapeutic target, accounting for the fact that effective treatments such as cognitive behavioural therapy for insomnia (CBTI) are substantially underused [21].
Although successful insomnia treatment might help to prevent other psychiatric disorders, particularly depression.

Despite these practical advantages, the concept of insomnia disorder has serious theoretical shortcomings challenging its usefulness:

**Unknown pathomechanisms probably unrelated to sleep physiology.** Even the most recent state-of-the-art reviews (e.g., [21,22]) acknowledge that the causes and pathophysiological mechanisms of insomnia are completely unknown. Of course, as with every psychiatric disorder, a complex biopsychosocial model of predisposing, precipitating and perpetuating factors for insomnia disorder has been proposed, dating back to the 1980s [23]. But it is also admitted that “decades of research in insomnia have failed to reveal circadian and homeostatic mechanisms as primary factors involved in the origin and pathophysiology” [24].

**Lack of systematic relationship between complaints and PSG parameters.** A recent approach to identify the potential biomarkers of insomnia disorder yielded six candidates [25]. Only one of these, the cyclic alternating pattern, was a PSG-derived sleep parameter, and only one single study suggested its usefulness. Instead, insomnia complaints are most strongly related to daytime symptoms such as fatigue and anxiety [26]. Moreover, objective and subjective sleep parameters show low to non-existing correlations, suggesting that they are essentially independent of each other [27]. This does not mean that there are no statistically significant differences at all between PSG sleep variables in insomnia disorder compared to controls (cf. [28]), but these are unsystematic, unspecific and demonstrate low effect sizes and are not of clinical value [29].

**Lack of insomnia disorder models.** Objective sleep parameters can be experimentally and reliably manipulated in animals and in humans. Sleep duration can be shortened by sleep deprivation, even selectively for REM and non-REM sleep. Sleep continuity can be challenged by acoustic or other stimuli. Sleep can be affected by various drugs or neurophysiological experiments. Hundreds of studies of this kind yielded a quite differentiated concept of sleep regulation (e.g., [8]) and numerous ideas about the functions of sleep (e.g., [5]). However, all these studies, their results and the concepts derived from them rely on objective measurements of sleep and, due to the lack of any systematic relationship to subjective complaints, cannot be transferred to insomnia disorder. Hence, we do not have any experimental model for insomnia disorder allowing for basic research in animals and humans.

**The insomnia disorder concept neglects concomitant objective alterations in sleep.** As a consequence of the independence of subjective perceptions of sleep from what we can measure, in clinical reality, we see people who maximally complain without any measurable alteration in sleep, people who do not complain at all but have seriously impaired sleep when we measure it, and all combinations one can imagine between these extremes. Figure 1 graphically represents this situation in a simplified way, assuming that both dimensions can be summarised as one continuous variable each. People above a certain threshold of subjective sleep disturbance are supposed to suffer from insomnia disorder. In clinical practice and in many studies, the degree of objective sleep disturbance in these people is entirely neglected, because PSG and PG studies are just not performed. Given an up to 25% prevalence of, e.g., sleep apnoea syndrome in psychiatric patients [30] and, vice versa, a prevalence of insomnia complaints in about 50% of those affected by sleep apnoea [31], this is of obvious concern.
The term “insomnia disorder” is misleading. Most people and probably many physicians do not know that subjective sleep-related complaints typical of insomnia disorder do not reliably indicate an actual reduction in the amount of sleep or a deterioration in sleep continuity with all its negative consequences. So, two largely independent dimensions are continuously being mixed up in the public perception of insomnia, and even in part in science, which leads to considerable confusion. A striking recent example of this confusion is the research on the dual orexin-receptor antagonist daridorexant. This drug has been approved for insomnia disorder in the US and Europe based on clinical studies, requesting that participants display a sleep onset latency of at least 30 min according to PSG recordings.

Negative physical consequences of insomnia disorder are probably unrelated to subjective sleep complaints. Insomnia disorder is supposed to have negative consequences on mental and physical health [21]. However, at least regarding its physical consequences, which may include diabetes, obesity, cardiovascular disease and alterations in immune functions, it is very likely that they are primarily linked to objective changes in sleep, rather than to subjective complaints. For example, glucose tolerance has been shown to be impaired in sleep apnoea and restless legs syndrome patients, who showed reduced sleep amounts and impaired sleep continuity PSG parameters compared to healthy subjects. In contrast to normal sleepers, about half of these patients were at increased risk for diabetes. Moreover, glucose tolerance was negatively correlated with the degree of sleep continuity disturbance [32]. In patients with insomnia, however, glucose tolerance was normal, as were their PSG results [32,33]. These findings do not indicate that insomnia disorder is without any consequences on physical health at all, but they suggest that these consequences might be limited to patients where an objective sleep pathology can be documented. This view is supported by a recent metaanalysis, which showed that in subjects with objective short sleep duration, diabetes and hypertension prevalence is increased, whether or not they report insomnia symptoms [34].

Figure 1. Subjective complaints versus objective sleep disturbances. For simplification, it is assumed that both dimensions can be represented by one continuous variable each. Zero means no complaint/disturbance at all; 100 denotes the maximum for both dimensions. The blue-shaded area arbitrarily defines insomnia disorder as being characterised by a degree of subjective complaints of 40 or higher. A total of 50 dots have been generated randomly; in the left upper and right lower areas, 4 dots have been removed to account for the fact that those extremes are rather unlikely. The figure illustrates that insomnia disorder can go along with almost every degree of objective sleep disturbance.
Moreover, PSG-measured latency to persistent sleep and wake time after sleep onset were chosen as primary endpoints [35]. Hence, the study population systematically differed from the target population and the empirically proven effects of the drug are at best loosely related to the symptoms of the disorder to be treated. Although two secondary subjective endpoints (subjective total sleep time and daytime functioning) improved as well, the approach taken is unique and somewhat disconcerting in light of all what is known about insomnia disorder.

To conclude, insomnia disorder is a typical mental disorder derived from a syndromal approach to subjective patients’ complaints without considering objective diagnostic data. This concept has proven successful in establishing pharmacological and psychotherapeutic approaches to these patient complaints [21]. However, it has been unsuccessful, and it is unlikely to be successful in the future, regarding the establishment of stringent aetiological and pathophysiological ideas. One important reason is that insomnia complaints are the consequence of a variety of distinct psychiatric and somatic disorders which are not stringently excluded when diagnosing insomnia disorder. No doubt, for example, insomnia complaints in a patient with obstructive sleep apnoea, or depression, or even both [30], have a different aetiological and pathophysiological background than in patients without these conditions. The second reason is that insomnia is seen as a sleep disorder, but biological sleep is not consequently taken into account. In the clinical context, by definition, sleep physiology plays no role at all. In research, on the other hand, sleep is studied physiologically, but the questions are addressed from the wrong perspective: one looks for biological parameters that explain, overall and specifically, the variable clinical picture of subjective complaints. Instead, it would probably be more promising to ask which physiological changes in sleep lead to which individual symptoms or sequelae (such as diabetes) independent of the syndromal entity.

Finally, it might be fair to ask whether insomnia disorder is a sleep disorder at all, or rather a change in perception and behaviour due to, for example, hyperarousal independent of the regulation of sleep and wakefulness [36].

5. Summary: The Future of Psychiatric Sleep Research

So far, the contribution of sleep research to the basic understanding of psychiatric disorders has been limited. Particularly, it has proven difficult to establish disorder-specific biomarkers from sleep recordings. However, at least in depression, there is preliminary evidence that elevated REM sleep pressure may predict the relapse and first occurrence of affective disorders. Here, again, prediction is not absolutely specific. Regarding insomnia disorder, measuring sleep has not proven helpful at all in unravelling its aetiology and pathophysiology so far. This is surprising at first glance but easy to understand if one accepts that the subjective perception of sleep and its quality is almost completely independent from biological sleep. However, the question is not who is right: the patient who complains or the scientist who measures. It just must be accepted that we are dealing with two independent dimensions. Ignoring this fact leads to confusion and failure in both basic research and clinical care.

Future sleep research should, rather than restrict itself to the search for specific sleep biomarkers in predefined complex disorders of perception and behaviour [37,38], use the opposite approach. Sleep assessed by PSG should be considered as an independent variable affecting perception and behaviour. This is a transdiagnostic approach starting from physiology rather than from subjective complaints. This idea is not new; in essence, it was already proposed by Tom Insel in 2010 [39], when he established the concept of Research Domain Criteria. Among those domains, one is called “arousal/regulatory” systems, representing the physiological regulation of wakefulness and sleep. Such an approach can not only lead to a better understanding of the relationship between objective measures and subjective approaches; it could also, in the long run, help to establish new, more valid diagnostic entities.
This approach would require that, in the future, patients with psychiatric disorders, on a large scale, not only be studied regarding their sleep-related complaints, but additionally be examined in the sleep laboratory. This would require a reduction in the time and effort necessary for both data collection and evaluation, and measurements would have to be carried out in the home environment, at least in many cases. The evaluation of recordings would have to be supported by modern evaluation algorithms, probably including machine learning. Of course, to establish the widespread use of physiological measurements of sleep in research and treatment of people with mental illness, some effort will still be needed, but the technical requirements for this will be available quite soon (e.g., [40,41]).

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The author declares no conflict of interest.

**References**


35. Markham, A. Daridorexant: First Approval. *Drugs* 2022, 82, 601–607. [CrossRef] [PubMed]


40. Penzel, T. Sleep scoring moving from visual scoring towards automated scoring. *Sleep 2022*, 10, 45. [CrossRef] [PubMed]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.