

<https://doi.org/10.1038/s42003-025-07794-6>

A continuous approach to explain insomnia and subjective-objective sleep discrepancy

Check for updates

Rubén Herzog¹, Flynn Crosbie², Anis Aloulou^{2,3}, Umaer Hanif^{1,2}, Mounir Chennaoui^{2,4}, Damien Léger^{2,3,6} & Thomas Andrillon^{1,5,6} ✉

Understanding insomnia is crucial for improving its diagnosis and treatment. However, many subjective complaints about insomnia do not align with objective measures of sleep quality, as is the case in subjective-objective sleep discrepancy (SOSD). We address this discrepancy by measuring sleep intrusions and instability in polysomnographic recordings from a large clinical database. Using machine learning, we develop personalized models to infer hypnodensities—a continuous and probabilistic measure of sleep dynamics—, and analyze them via information theory to measure intrusions and instability in a principled way. We find that insomnia with SOSD involves sleep intrusions during intra-sleep wakefulness, while insomnia without SOSD shows wake intrusions during sleep, indicating distinct etiologies. By mapping these metrics to standard sleep features, we provide a continuous and interpretable framework for measuring sleep quality. This approach integrates and values subjective insomnia complaints with physiological data for a more accurate view of sleep quality and its disorders.

Sleeping is essential to our survival¹. However, between 10 and 20% of adults living in industrialized countries struggle to achieve a good night's sleep and complain of chronic insomnia². Poor sleep and insomnia take a significant toll on general and mental health (e.g., increase in cardiovascular risks, anxiety, depression, accidents, etc)^{3–6}. Because of their high prevalence and far-reaching consequences, tracking, managing, and curing sleep difficulties are paramount for public health and quality of life^{7,8}. However, the diagnosis and treatment of sleep disturbances can be complex, notably because of a lack of robust biomarkers for sleep quality. This difficulty is partly due to the frequent discrepancy between individuals' subjective reports regarding their sleep and the findings of objective standard sleep exams, a phenomenon referred to as sleep state misperception (SSM) or subjective-objective sleep discrepancy (SOSD)^{9,10}.

Quantifying sleep is complex and can be done in various ways, such as through subjective reports (e.g., questionnaires, scales, sleep logs, etc), measurements of body activity (e.g., using actigraphy, a wrist-accelerometer), or measurements of brain activity. The latter is performed with the technique of polysomnography (PSG), which consists of the continuous recording of brain (electroencephalography, EEG), ocular (electrooculography, EOG), muscular (electromyography, EMG) activity¹¹. PSGs

are the gold standard of sleep research and serve as the basis for the definition of sleep and its sub-stages^{12,13}.

PSGs are required to confirm the diagnosis of all sleep disorders, with the notable exception of insomnia^{14,15}. This is perhaps because, in up to 50% of insomnia complaints, the visual inspection of PSG recordings significantly diverges from self-reported estimates of sleep quantity or quality⁹. Consequently, chronic insomnia is defined on subjective reports alone (namely reported difficulties in initiating or maintaining sleep and/or a chronic sensation of non-restorative sleep¹⁶). However, pharmacological treatments for insomnia still need to be evaluated based on criteria derived from PSG recordings¹⁴. For example, the American Food and Drug Agency requires positive effects on PSG-derived sleep metrics to authorize new drugs, and the European Medicines Agency recommends using PSGs in clinical trials^{17,18}.

Beyond the specific case of insomnia, SOSD questions the overlap between the objective and subjective estimates of sleep^{10,19,20}. SOSD is present in all sleep disorders^{21,22} and, although to a lesser degree, this is a phenomenon also recognized in good sleepers¹⁰ and 20% of hypersomnia complaints. SOSD can be bidirectional and reflect the feeling of being awake while PSG recordings indicate sleep, but also the feeling of being asleep while

¹Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, Paris, France. ²Université Paris Cité, VIFASOM (Vigilance Fatigue Sommeil et Santé publique), Paris, France. ³APHP, Hôtel-Dieu, Centre du sommeil et de la Vigilance, Paris, France. ⁴Institut de recherche biomédicale des armées (IRBA), Brétigny-sur-Orge Paris, France. ⁵Monash Centre for Consciousness & Contemplative Studies, Monash University, Melbourne, Australia. ⁶These authors jointly supervised for this work: Damien Léger, Thomas Andrillon. ✉e-mail: thomas.andrillon@icm-institute.org

PSG recordings indicate wakefulness¹⁰. In good sleepers, a recent study stressed the lack of correlation between classical neurophysiological indexes of sleep depth and subjective sleep depth²³. These results deeply question some fundamentals of sleep science and ask for renewed efforts to identify robust biomarkers of sleep quality in order to better understand what makes a good night's sleep.

Sleep is generally associated with a loss of responsiveness and awareness to the external world, which are accompanied by changes in brain activity with a shift to high-amplitude low-frequency oscillations²⁰. Based on this relationship, the current consensus is that the detection of certain EEG hallmarks (e.g. slow waves, sleep spindles) is enough to identify the state of sleep and to infer sleep depth. However, recent discoveries show that this characterisation of sleep and sleep depth is too coarse²³, does not capture fine transitions at sleep onset or between sleep sub-stages²⁴, and most importantly does not always align with subjective feelings²⁵. This could be first because standard PSGs rely on a small number of EEG electrodes (typically just 3), providing a partial reading of sleep state and its local regulation^{26,27}. This is important because local modulations of sleep depth can account for the feeling of being awake while asleep²³, the occurrence of dreaming²⁸, the sensitivity to sensory inputs during sleep²⁹, and sleep disturbances such as insomnia³⁰. Second, relying on the visual inspection of sleep recordings could mask important signs of sleep disturbances that are not visible to the naked eye^{31,32}. For example, fast oscillations are difficult to track visually, and are even filtered out when specialists inspect PSGs, despite the fact that several studies have linked these fast oscillations to S OSD^{10,31,33}. A recent examination of the topography of PSG has associated diffuse cortical hyperactivations with S OSD, identifying a decrease in the delta/beta ratio as a major correlate³⁴. However, for typical lower-density clinical PSG recordings, it is still difficult to track the subtleties that could explain how a subjective complaint of insomnia can coexist with what appears to be, at the surface, normal sleep on the PSG³¹.

New methods in sleep medicine could help reduce the gap between objective and subjective assessment of sleep and uncover the root causes of poor sleep quality. With increasing computational power, it is now easy to extract hundreds or thousands of different features even in large datasets³⁵, considerably enriching the quantity of information derived from a single PSG recording. This approach corroborates and nuances the standard classification of sleep in a data-driven way³². Also, advances in machine learning further enhance this potential for identifying the optimal set of features for classifying a specific sleep disorder, uncovering new biomarkers for diagnosis and elucidating new mechanisms for more targeted and effective treatments. Furthermore, recent studies on automated sleep stage classification illustrate these opportunities³⁶. Visual scoring of sleep stages, as rapid eye movement (REM) or non-REM (NREM) sleep, crucial yet time-consuming, achieves around 80% interscorer agreement³⁷, while machine-learning techniques are capable of reaching human performance levels in healthy and clinical groups^{38–40}. Beyond practical classification, these algorithms offer a probabilistic perspective on sleep by computing the probabilities of the five standard sleep stages, known as hypn densities, which provide a finer and continuous description of sleep, that could aid in diagnosis^{38–40}.

Here, we set out to leverage innovations in big data and machine learning for the detection of insomnia and S OSD. We retrospectively analyzed a large cohort of patients diagnosed with chronic insomnia based on subjective complaints¹⁶. We sought to identify the neural signatures that distinguish between insomnia with and without S OSD using PSG recordings. While prior research has highlighted differences in various markers of sleep quantity and quality¹⁰, accurately identifying insomnia with S OSD based solely on PSG data remains a challenging endeavor. All these patients underwent a PSG and we found evidence for a significant S OSD in a subset of these patients (S OSD+ group, 24% of total) and no evidence of a significant S OSD in the rest (S OSD– group) (see³¹ and Methods for details) although the exact definition of what constitutes S OSD varies from one study to another⁴¹. We compared these two insomnia groups with healthy volunteers (good sleepers, the GS group). From PSG recordings, we

extracted a set of highly discriminative and minimally redundant features⁴², which we used to train and validate a personalized model of sleep architecture. From this personalized model, we extracted hypn densities to train a cross-subject algorithm for the detection of insomnia with or without S OSD, reaching an overall accuracy of $0.77 \pm 0.017\%$. We finally interpreted the internal functioning of the algorithm (i.e., opening the black box) to identify the neurophysiological signatures of insomnia and revealed evidence for sleep/wake mixing in both sleep and wakefulness for the S OSD– and S OSD+ groups. The S OSD+ group was further characterized by higher probabilities of sleep within wakefulness and higher wake instability, suggesting that S OSD stems from a perturbation of both sleep and wakefulness. Finally, our approach was also sensitive to severity of objective impairment of polysomnography in insomnia as we could accurately predict markers of sleep quality (total sleep time [TST], sleep onset latency [SOL], wake after sleep onset [WASO] and sleep efficiency [SE]) from PSG recordings. Overall, this work shows that finer analyses of PSG recordings can close the gap between subjective complaints of insomnia and objective quantifications of sleep quality.

Results

Hypn densities: leveraging a probabilistic approach to sleep staging

We examined a large database of PSG recordings ($n = 927$) to identify sleep quality indicators that vary between the three groups of sleepers: Good sleepers (GS, $n = 104$), patients with insomnia without sleep state misperception (S OSD–, $n = 624$) and patients with sleep misperception (S OSD+, $n = 199$) (Fig. 1A). The clinical dataset utilized was diverse in terms of recording devices, electrode layout, and sampling rates, making direct comparisons challenging. To overcome this issue, we developed an analytical framework that leverages the estimation of hypn densities (HD)³⁸. Expanding on the standard discrete sleep staging, HD provides a nuanced assessment of the probability of each sleep stage (wake, N1, N2, N3, and REM) within a given epoch. From the HD computed on each epoch, we measured the level of stage mixing and epoch instability (see *Methods: Intrusions and instability*), and used these measures as biomarkers to predict the clinical group and standard sleep quality metrics at the individual level (Fig. 1B). Sleep and wake intrusions were defined respectively as the probability of wakefulness during sleep epochs and the probability of sleep during wake epochs.

We first devised a computationally-light approach to compute HD from PSG recordings. We employed *catch22*⁴², a robust feature extraction algorithm, to extract features from each EEG and EOG channel, enabling us to project each 30-second epoch into a multidimensional space. Then, by leveraging the sleep scoring done by experts (see *Methods: Hypnograms*), we trained a machine learning (ML) algorithm to predict the hypnogram based on these features, yielding a HD for each epoch (see *Methods: Data harmonization by hypn density estimation*). Finally, we used tools from information theory, namely the entropy and the Kullback-Leibler divergence (D_{KL}) to quantify the level of intrusion and instability of each HD, respectively (Fig. 1C; see *Methods: Intrusions and instability* for a detailed definition and interpretation of these metrics).

Increased intrusions in insomnia: sleep intrusion in S OSD+ and wake intrusion in S OSD–

We started by comparing our predicted sleep staging with the sleep staging obtained with human sleep experts and selecting only subjects where our pipeline yielded hypnogram prediction accuracy larger than 0.4 (Supplementary Fig. 1A, B. 115 subjects were discarded, whose data was not included in this work). As a first sanity check, we confirmed that, on average, the maximum of each HD matched its corresponding epoch stage (Fig. 2A), with the exception of N1, which was also comparatively poorly predicted by the algorithm (Supplementary Fig. 1C, D). For this reason, N1 was excluded from the expert stages to be analyzed, but it was nevertheless considered in the computation of HD, entropy and DKL for completeness. In the

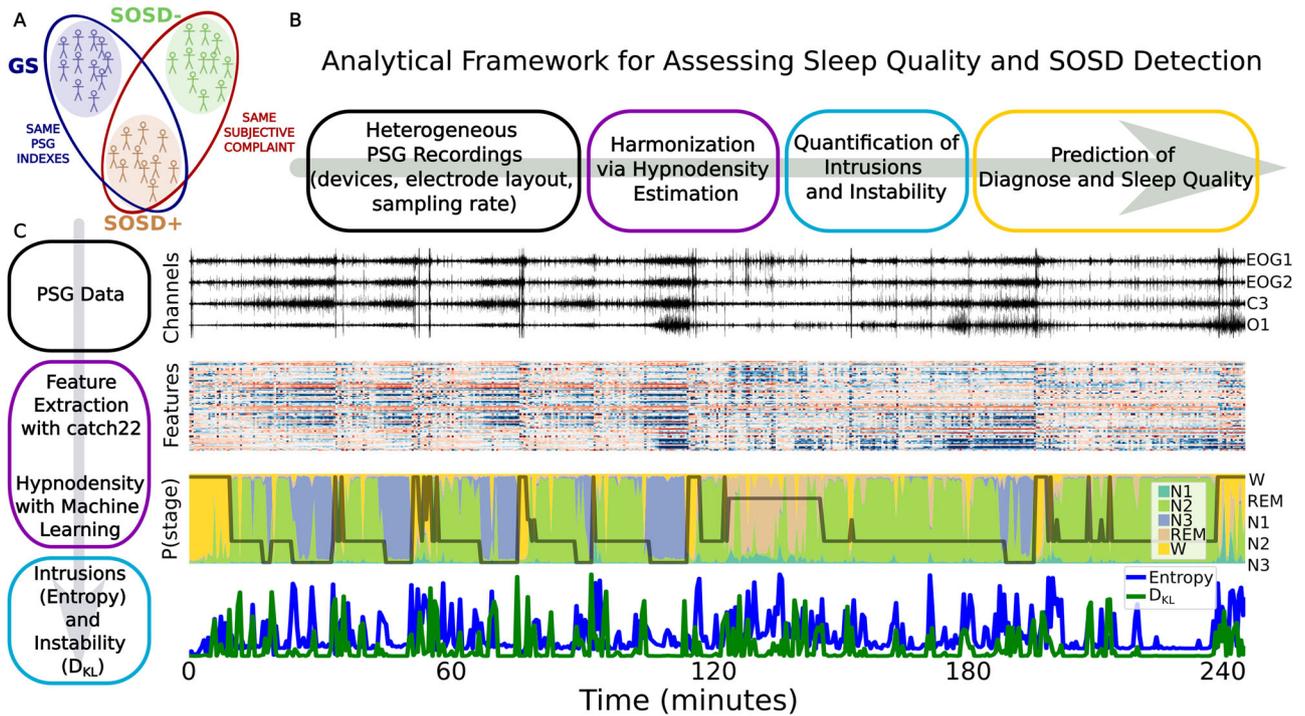
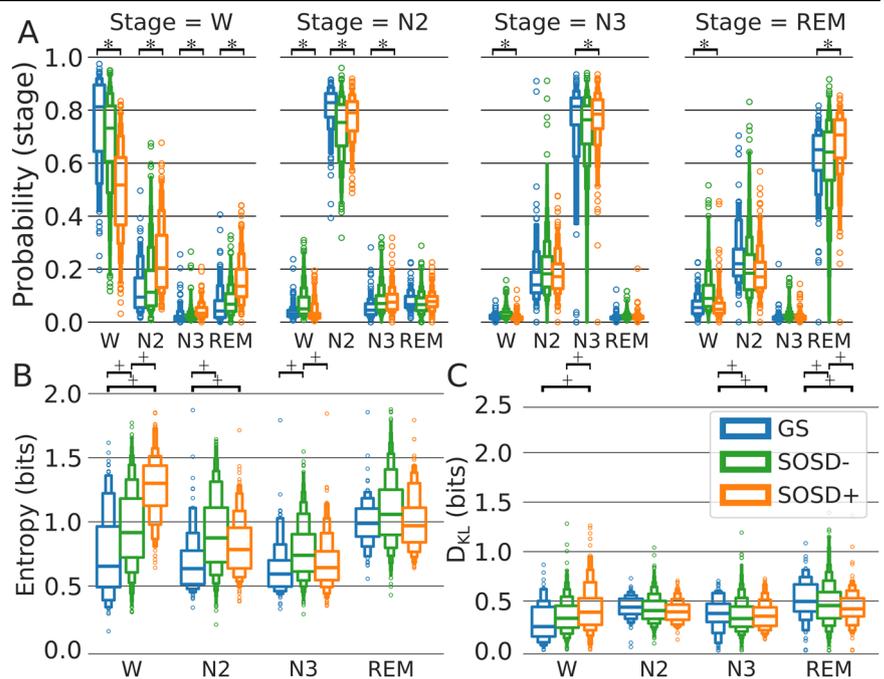


Fig. 1 | Analytical pipeline for characterizing sleep quality and addressing subjective-objective sleep discrepancy (SOSD). **A** We analyzed PSG data from a population of Good Sleepers (GS), patients with insomnia without SOSD (SOSD-) and with SOSD (SOSD+). **B** Heterogeneous PSG recordings were harmonized via the estimation of hypnodensities, which enabled the quantification of the level of intrusion and instability of sleep dynamics. Finally, intrusions and instability were used as markers for predicting individual diagnoses and sleep quality measures.

C PSG recordings (example of the first 4 h of sleep of a GS subject) were scored by experts in 30 s epochs (hypnogram) and then a set of features (*catch22*) were extracted to characterize each epoch in a multidimensional space. A machine learning algorithm was trained to predict the hypnogram. This also yielded, for each epoch, the probability of each stage, i.e., the hypnodensity (colors). Finally, the hypnodensities were analyzed with tools from information theory to measure the level of intrusions (entropy) and instability (D_{KL}) of each epoch.

Fig. 2 | Intrusions and instability on good sleepers and in insomnia with and without SOSD.

A Boxplots with average hypnodensities inferred for each sleep stage scored by experts and for each sleep group (colors). Black horizontal lines in the top of the panels denote significant differences between groups (* represents Bonferroni corrected Kruskal-Wallis $p < 0.001$). **B, C** The respective average level of intrusion (entropy) and instability (D_{KL}) for each sleep stage scored by experts and sleep group, respectively. Black horizontal lines on top of the panels denote significant pairwise differences (+ represents Bonferroni corrected Wilcoxon rank sums $p < 0.001$). SOSD-: $n = 624$, SOSD+ : $n = 199$ and GS: $n = 104$; Each sample corresponds to a different participant.



following, any mention of sleep stages will refer to those scored by experts, while mentions to stage probabilities or intrusions refer to the probabilities obtained from the hypnodensities.

When examining group differences, we found a significant reduction (Bonferroni corrected Wilcoxon-rank sum $p < 0.001$) of the wake probability during expert-scored wakefulness for SODS- (0.70 ± 0.16) and even less for SODS+ (0.49 ± 0.17), compared to GS (0.75 ± 0.18) (see Supplementary Table 1 for HD details). While SODS- exhibited a significantly larger probability of wakefulness during sleep (NREM + REM), SODS+ did not (GS: 0.091 ± 0.14 ; SODS-: 0.12 ± 0.12 ; SODS+ : 0.067 ± 0.080); Bonferroni corrected Wilcoxon-rank sum GS vs SODS- $p < 0.001$, GS vs SODS+ $p > 0.01$). In fact, SODS- exhibited larger wake intrusions (defined here as the wake probability in sleep epochs) than SODS+ in all sleep stages. Thus, our results show that both SODS- and SODS+ insomnia sub-types are associated with sleep intrusions during wakefulness, but in the case of SODS+ sleep intrusions are not accompanied by wake intrusions during sleep. These results are in line with the apparent normal sleep characteristic of SODS+ (see Supplementary Table 2–5 for details on sleep microstructure of the groups).

To quantify and compare the levels of intrusion between sleep group and sleep stages, we computed the entropy of each HD (see *Methods: Intrusions and instability*). Entropy is a function of the full HD that measures the proximity to a uniform distribution, thus providing more information than just measuring the highest probability on the HD. As high HD entropy indicates a high level of intrusion (proximity to a uniform distribution), high entropy during wakefulness implies sleep intrusions, while high entropy during sleep implies wakefulness or intrusions from other sleep stages. Consistent with the previous results, we found a significant increase in HD entropy during wakefulness in both insomnia groups (with and without SODS) compared to GS and also in SODS+ compared to SODS- (Fig. 2B, GS: 0.74 ± 0.34 bits; SODS-: 0.95 ± 0.30 bits; SODS+ : 1.27 ± 0.24 bits; Bonferroni corrected Wilcoxon ranksum $p < 0.001$ for all pairwise comparisons). Note that these differences held even after controlling by the stage normalized frequency (i.e., the number of epochs of each stage divided by the total number of epochs; see Supplementary Table 6). Also, SODS- had a significantly larger entropy during N2 and N3 than GS, and a significantly larger entropy in N3 than SODS+ (Bonferroni corrected Wilcoxon ranksum $p < 0.001$). This increase in entropy in NREM sleep in the SODS- group can be largely attributed to an increase in wake probabilities. In contrast, while SODS+ also has a significantly larger entropy in N2 than GS, this increase was not related to an increase in wake intrusions, but rather to an increase in the probability of N3 (so deeper sleep). See Supplementary Table 7 for details.

To further characterize the sleep dynamics, we quantified the level of instability between two consecutive HDs within the same sleep stage by the D_{KL} . A low D_{KL} means that both epochs are very similar in terms of their HDs (i.e. high epoch-to-epoch stability), while a high value means that two consecutive epochs largely differ. Note that this is different from just computing the difference of entropies, as two HDs could be different but still have the same entropy (e.g. same probabilities but assigned to different stages). We found that SODS+ had larger wake instability than GS and SODS- (Fig. 2C, Bonferroni corrected Wilcoxon ranksum $p < 0.001$). These differences held after controlling for sleep stage normalized frequency, while SODS- did not significantly differ from GS. No differences were found for N2. Both SODS- and SODS+ had less instability than GS for N3, with no significant differences between them. Finally, REM was more unstable for GS than in SODS- and SODS+, with SODS+ exhibiting less instability than SODS-. However, after controlling by the stage normalized frequency, the difference between SODS+ and SODS- on REM vanished (see Supplementary Table 8). To summarize, reduced instability in N3 and REM is a common factor in the perception of poor sleep quality in SODS- and SODS+. Yet, only SODS+ shows an increased instability in wakefulness compared to GS. Again, these results are consistent with the condition of SODS+ as a disruption of intra-sleep wakefulness rather than sleep itself.

Sleep intrusions and wake instability predict diagnosis and sleep quality

To further confirm that intrusions and instability were reliable markers of sleep groups, we used these features as predictors in a machine learning classifier (XGBoost, see *Methods: Matching learning classifier and regression*). We performed two different types of classifications: a 3-class (GS vs SODS- vs SODS+) classification and a within-insomnia 2-class (SODS- vs. SODS+) classification. In detail, we calculated the average and standard deviation of intrusions and instability for each sleep stage, yielding a total of 16 features (4 features per sleep stage, 4 sleep stages analyzed). We then applied feature selection to select the smallest set of features that maximizes classification performance. First, we sequentially included the features in the classifier according to their maximum relevance minimum redundancy score (MRMR, see *Methods: Feature selection*). To avoid biases given by size imbalances, we took 30 random subsamples of a size equivalent to the smallest sample size (GS in the 3 class problem and SODS+ in the binary problem, see *Methods: Matching samples by size, sex, age and BMI*) and ran a repeated 5-fold cross-validated classifier. In both cases, the 3 first features were intrusions in wakefulness, instability in wakefulness and the standard deviation of instability in REM. We found better performance for the 2-class than for the 3-class classifiers (Fig. 3A), but, in both, the performance was saturated with less than 10 features. We examined the confusion matrices (Fig. 3B, C) associated with the points before performance saturation (vertical lines in Fig. 3A; 9 and 4 features for the 3-class and 2-class classifiers, with average AUC of 0.82 ± 0.016 and 0.86 ± 0.013 , respectively) and confirmed a strong diagonal in both cases. We repeated the analysis for the 2-class classifier using the best 4 features in subsamples matched for age, sex and body-mass index and found a small performance decrease (Supplementary Fig. 2, average AUC of 0.83 ± 0.014 and 0.86 ± 0.013 for matched and unmatched samples, respectively; Wilcoxon ranksum $p < 0.001$). Finally, we repeated the analysis using hypnodensities derived from the U-Sleep algorithm⁴⁰, which is a state-of-the-art algorithm for automated sleep staging and hypnodensities estimation. This additional analysis was also designed to test if an algorithm trained on a different dataset would extract hypnodensities that allow the classification of the three groups even when the model is not trained on the PSG recordings themselves. Following the same feature selection and inclusion procedure, the hypnodensity derived from this external model yielded lower but above chance classification performance (Supplementary Fig. 3). These results are consistent with the previous section and demonstrate that intrusions and instability in wakefulness not only characterize the differences between sleep groups, but they can also be used to distinguish insomnia subtypes at the individual level.

However, the classifiers did not reach a perfect performance and to better understand cases of misclassification, we related the subject prediction accuracy of the 2-class classifier with the sleep metrics used for diagnosis (Fig. 3D, E; see *Methods: Participants and diagnose*). Subject prediction accuracy was obtained by running 1000 iterations of the repeated cross-validated classifier and counting how many times a subject was correctly classified in the test sample. This value is 1 when the subject is correctly classified in all the test samples and 0 for the opposite case. We found that most of the misclassifications occur for patients close to the cut-offs used for diagnosis (see *Methods: Objective insomnia*), suggesting that the thresholds used for sub-typing of insomnia could be hiding a continuum within insomnia.

To further explore the idea that intrusions and instability could index sleep quality in a continuous way, we used the same features to predict the values of the sleep quality metrics in insomnia patients. TST, SOL, WASO and SE are critical metrics for the diagnosis and evaluation of insomnia severity¹⁶. To do so, we used 120 iterations of a repeated 5-fold cross-validated ML regression algorithm. As for the previous classification problem, we used the MRMR scores to include features sequentially as predictors of the sleep metrics and measured the regression performance using the RMSE (Fig. 4A). In all cases we found that adding more features improved the performance, with no evident saturation point this time. All

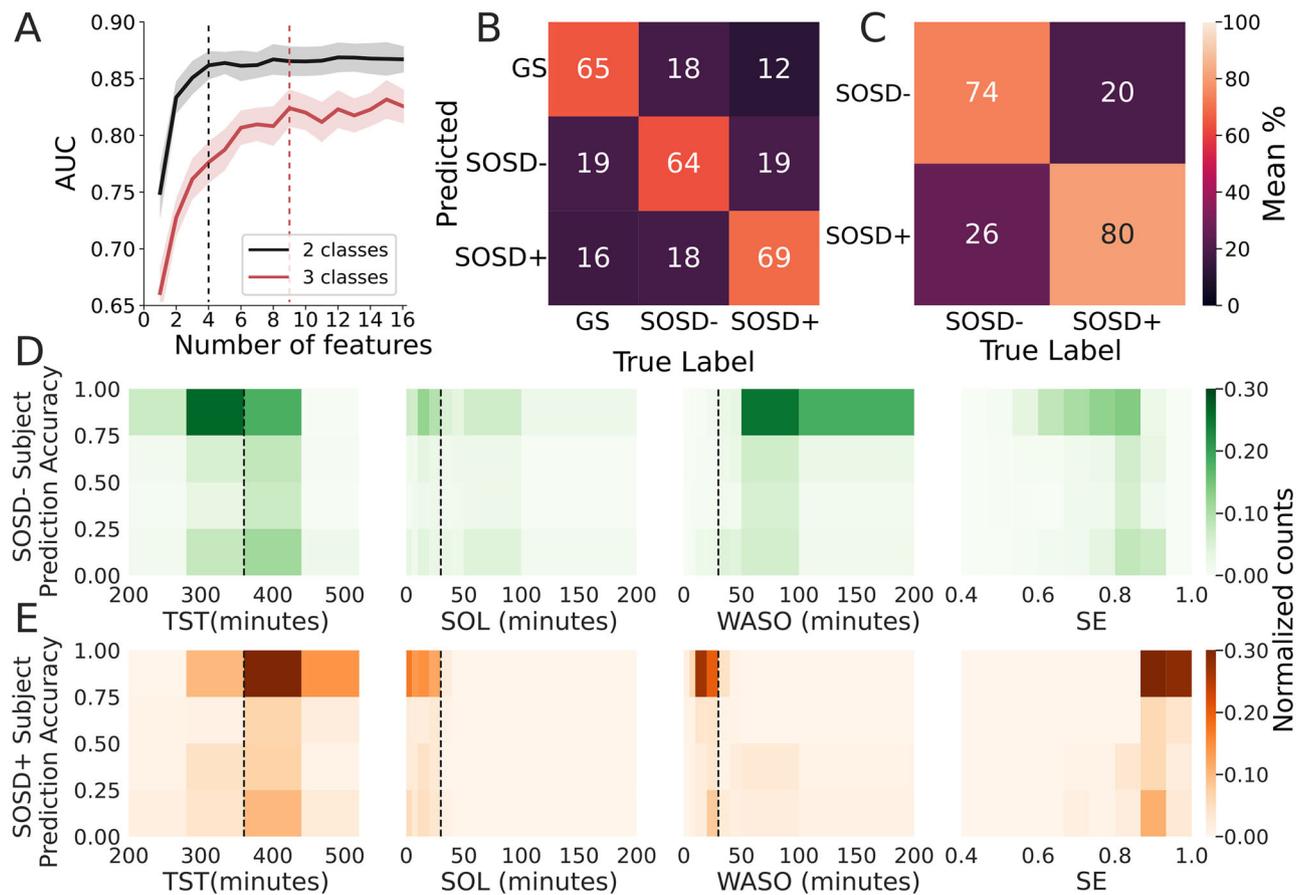


Fig. 3 | Sleep intrusions and wake instability predict sleep groups. **A** Classifier average performance as a function of the number of features included as predictors for binary (black, only SODS- and SODS+) and 3 class problems (red, GS, SODS- and SODS+) using size-balanced subsamples of data. For the 3-class problem all random subsamples of SODS- and SODS+ comprised 104 subjects, while for the binary problem random SODS- subsamples of 199 subjects were taken, so all the sleep groups had the same sample size. Solid line and shaded area correspond to the

average and 1 standard deviation, respectively, of 60 iterations of a repeated 5-fold cross-validated classifier, using 20 random subsamples each time. **B, C** Average confusion matrix for the 3 and 2-class problem, corresponding to the vertical lines in **(A)**. **D, E** 2D Histograms of the class prediction accuracy for each subject in the binary problem and sleep quality metrics for SODS- and SODS+, respectively. Color scale represents the percentage of the population and vertical dashed line represents the cut-off for the diagnosis associated with each sleep metric.

the predictions were significantly correlated with the veridical variables obtained from PSG recordings (Fig. 4B). We obtained the best correlations for WASO followed by SE, TST, and SOL. Predictions were also better for SODS- than for SODS+. The average intrusions and instability in wakefulness had the largest correlation with WASO and SE (Fig. 4C), with a negative correlation for the former and a positive for the latter. Accordingly, our results suggest that the level of intrusions and instability in wakefulness can not only discriminate between sleep groups, but also predict the severity of objective markers for the diagnosis of insomnia.

Sleep intrusions correlate with low-frequency EEG activity

Finally, we aimed to understand intrusions and instability in terms of PSG macro and microstructural features (transitions between sleep stages and spectral analysis, respectively). Following previous results on dissociated states in sleep and wakefulness^{20,26} in particular in insomnia and SODS^{23,30}, we expected that sleep intrusions during wakefulness would be characterized by increased EEG power in low frequencies, while for intrusions during sleep we expected an increase in the power of high frequencies. Accordingly, we found that intrusions in wakefulness were positively correlated with the power in delta and theta bands, while for sleep stages intrusions were positively correlated with the power in high frequencies, specifically in the beta and gamma bands (Fig. 5A). This effect was present regardless of the EEG electrode position (Supplementary Fig. 4). Notably, in N3, intrusions were inversely correlated with the power in the delta band, which is often

used as a proxy for sleep depth⁴³. The same trend was observed when splitting by sleep groups (Fig. 5D), with subtle differences within each sleep stage and frequency band. For example, intrusions for SODS- and SODS+ were more correlated than GS for low-frequency bands (delta) in wake and for high-frequency bands (gamma) in N3. Note that our procedure to infer HD did not involve measuring power in these bands, demonstrating that despite the novelty of our approach it is highly consistent with previous results on dissociated sleep states^{10,44,45}.

Instability followed a similar but lesser tendency when correlated with power in frequency bands (Fig. 5B). A more detailed analysis revealed that instability within sleep stages correlated more with increases in the power of high frequencies in two consecutive epochs (Fig. 5C), especially in N3, which was also present when splitting by sleep groups (Fig. 5E). Moreover, instability in wakefulness was negatively and significantly correlated (Bonferroni corrected Pearson $p < 0.001$) with the probability of wake self-transition, an observation that held when splitting by sleep groups (Fig. 5F).

In conclusion, our results show that insomnia and insomnia subtypes can be detected by an automated analysis of hypnodensities. Furthermore, our results show evidence of wake intrusions during sleep for SODS- patients and of sleep intrusions during wake both SODS- and SODS+ patients. This is consistent with previous findings showing that sleep intrusions during wakefulness are associated with increases in EEG low frequency power²⁰ and that wake intrusions during sleep are associated with increases in high-frequency EEG power. Thus, our approach does not only

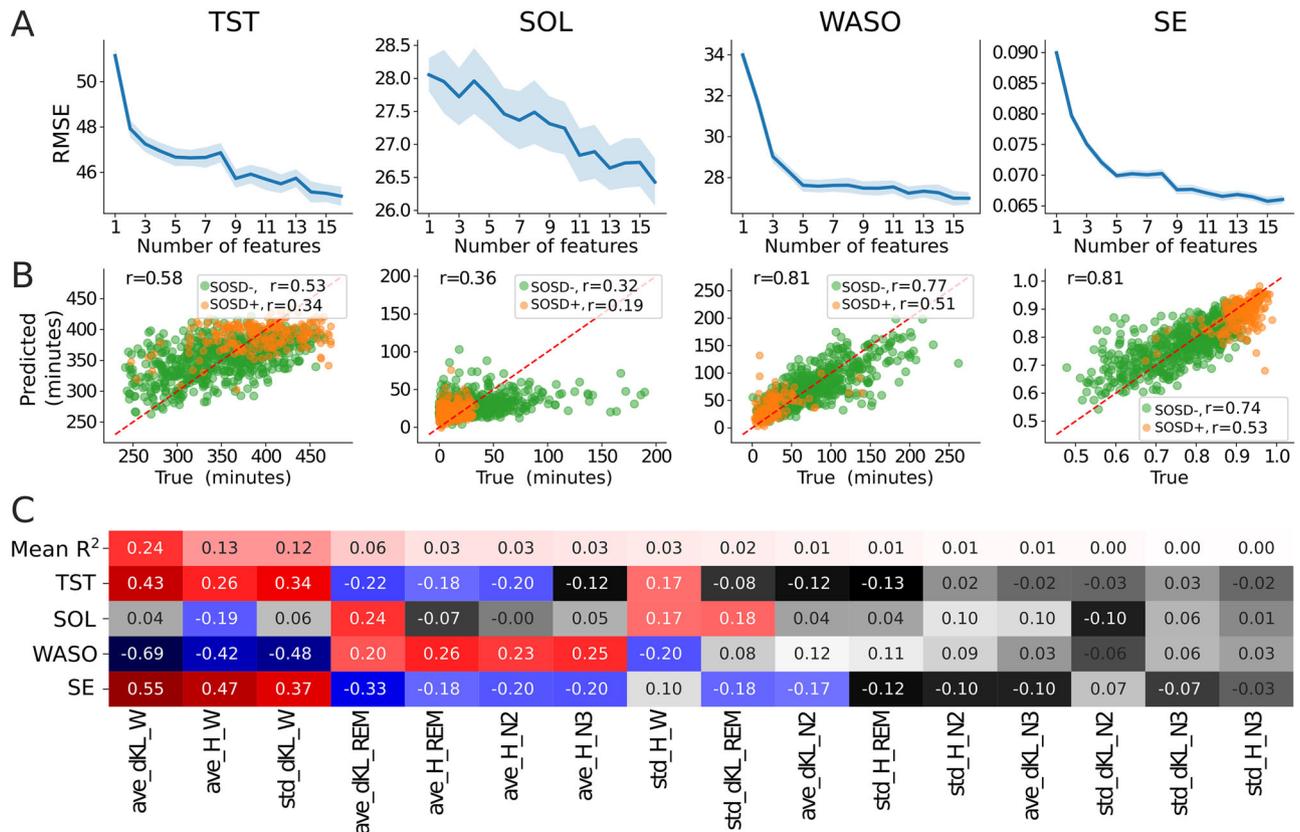


Fig. 4 | Intrusions and instability predict sleep quality. **A** Root mean squared error (RMSE) for a ML regression as a function of the number of features included as predictors for each of the 4 sleep parameters, using only SOSD+ and SOSD-. Lines and shaded areas are the average and standard deviation, respectively, out of 120 iterations of a repeated 5-fold cross-validated ML regression algorithm. **B** Average predicted vs true sleep parameter values corresponding to using all the 16 features. Subject averages were obtained from the predictions in the test samples. Colors denote sleep groups and the diagonal dashed red line is the identity line. All metrics significantly correlated with their prediction ($p < 0.001$), and the correlation

followed WASO > SE > TST > SOL. Predictions were better for SOSD- than for SOSD+. **C** Spearman correlation matrix between the features and the sleep quality parameters. Only colored squares were significant (Bonferroni corrected p -value < 0.001). Features are sorted in descending order following the average R² (first row of the matrix). Features are encoded as 'ave' for average, 'H' and 'dKL' for entropy and D_{KL}, respectively, and the last word corresponds to the sleep stage. For example, ave_H_W is the average entropy of wakefulness. Note that average intrusions and instability in wakefulness show the largest average R² and the largest correlations with WASO and SE.

lead to good classification performance but also allows us to interpret the underlying features, shedding a new light on the pathophysiology of insomnia.

Discussion

We aimed to identify neural signatures differentiating insomnia with and without SOSD using PSG recordings. Previous research demonstrated differences in various markers of sleep quantity and quality but the identification of insomnia with SOSD on PSG alone proved difficult¹⁰, which potentially stresses the need for large sample sizes and finer methods. In particular, PSG datasets can include large differences both within and between datasets, such as variations in the equipment used, the number or location of EEG electrodes and the recording or preprocessing routines. We addressed this issue by implementing data harmonization through the extraction of hypnodensities. Our aim here was to mitigate dataset-specific idiosyncrasies and enhance comparability across studies and experimental setups. From a more fundamental perspective, hypnodensities also capture the fact that sleep is not composed of discrete, mutually exclusive states^{20,24} but mixed sleep states are both frequent⁴⁶⁻⁴⁸ and informative of sleep quality^{23,31}. By leveraging information theory and ML techniques to analyze hypnodensities, we provided a nuanced understanding of insomnia, potentially applicable to other sleep disorders⁴⁹. Our findings revealed significant differences in the physiological profile of SOSD+ and SOSD-, despite similar subjective complaints, reinforcing the view of SOSD+ as a distinct insomnia sub-type⁵⁰. Both conditions exhibited increased sleep

intrusions during intra-sleep wakefulness, with a higher magnitude observed in SOSD+. In contrast, only SOSD- was characterized by significant wake intrusions during sleep. This distinction aligns with the "era of intrusions" described by Stephan and Siclari and give support to the hypothesis that SOSD is actually a mismeasurement rather than a misperception¹⁰. Despite the lack of spatial resolution of our work, our results are compatible with the recent finding of diffuse cortical hyperarousal as a marker for SOSD³⁴, which could be also interpreted as sleep/wake mixing. Our results emphasize the necessity to shift from binary diagnoses (as SOSD- vs SOSD+) to a graded diagnostic framework that better embraces the variability within sleep and between individuals²⁴.

Our results highlight the contribution of both sleep and wakefulness to insomnia, in line with the hyperarousal model of insomnia⁵¹. According to this model, insomnia would result from heightened levels of physiological and psychological arousal throughout the day and night. In support of the hyperarousal model, individuals with insomnia show increased cardiac activity both before and during sleep, elevated temperature during sleep, increased metabolism during wakefulness, and increased levels of cortisol during wakefulness (see refs. 52,53 for reviews). In our analyses, we found markers of wake intrusions during sleep for the SOSD- group (higher entropy and wake probabilities) but not for the SOSD+ group (Fig. 2), while intrusions and instability in wakefulness were good predictors of some of the commonly used sleep quality metrics (SE, TST and WASO). Correlations with spectral features indicate that this higher entropy is positively correlated with increase in fast wake-like frequencies (beta, gamma; Fig. 5). We

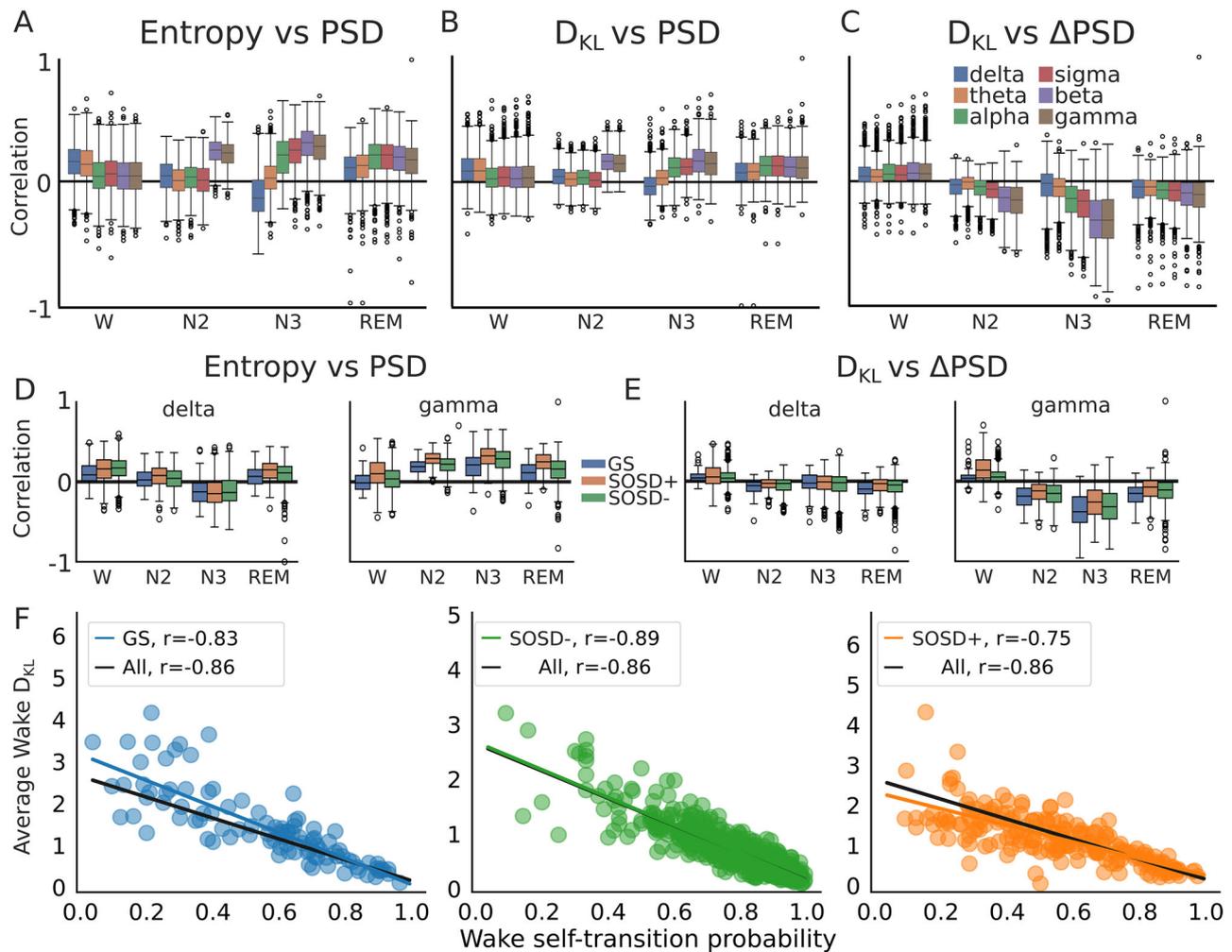


Fig. 5 | Interpretation of intrusions and instability in terms of macro and microstructural PSG features. **A** Boxplots with the population correlations between entropy and spectral power in 6 EEG frequency bands for electrode C3. **B** Same as (A), but for the D_{KL} . **C** Same as (B), but using the difference between power in two consecutive epochs within the same sleep stage. Δ PSD corresponds to $PSD(t) - PSD(t + 1)$, where t corresponds to a specific epoch. **D** Same as (A), but

split by sleep groups, and showing only delta and gamma bands. **E** Same as (D), but for the D_{KL} vs Δ PSD. **F** Correlation between the average D_{KL} in wakefulness and the wake self-transition probability, for the 3 sleep groups (colors). Black line is the correlation for the whole population and colored lines are the correlation for each sleep group.

thus found evidence of hyperarousal during sleep but only for SOSD– patients. SOSD+ patients showed a different pattern of results with an increase in entropy and Kullback-Leibler divergence, depicting intrusions and instability respectively, mostly during wakefulness (Fig. 3). Intrusions during wakefulness were characterized by increased low-frequency EEG activity (Fig. 5). As in the SOSD+ group, SOSD– patients also showed intrusions during NREM sleep, this time correlating with fast-frequency EEG activity. First, these results suggest that the hyperarousal model may not fully apply to SOSD+, for which wakefulness shows signs of hypoarousal rather than hyperarousal. Second, SOSD– and SOSD+ could represent different extremes of disrupted arousability leading to a subjective feeling of a lack of sleep: the former by hindering the ability to fall asleep and the latter by compromising wakefulness quality. This novel insight could not only inform diagnosis but also treatments for insomnia based on these different processes, potentially shifting the clinical management of SOSD+ from pharmacological treatments seeking to improve sleep to therapeutic approaches seeking to improve daytime vigilance such as light therapy or physical exercise¹⁵. Indeed, the current reference treatment for insomnia is cognitive-behavioral therapy for insomnia (CBTi), whose effectiveness is recognized regardless of the type of chronic insomnia⁵¹. We propose that the notions of “intrusions”, “stability” and “mixing” of sleep states represent

useful concepts to explain the correspondence between what is measured by the PSG and what is perceived by individuals, to enable certain patients to better understand the complexity of brain states during sleep. These notions could help practitioners explain to patients the discrepancy between a normal PSG outcome and their insomnia complaint. Prospective studies could also leverage this new approach to personalise treatments by investigating if metrics derived from PSG recordings can predict treatment outcomes.

We confirm our previous findings³¹ that insomnia can be detected from PSG recordings with high performance (AUC: 0.86; Fig. 3) and extend them by showing that the same PSG recordings can be used for the subphenotyping of insomnia and the distinction between SOSD+ and SOSD– (AUC: 0.82; Fig. 3). Furthermore, we highlight the limitations of traditional discrete classifications of SOSD– and SOSD+, suggesting that a continuum model more accurately captures the complexities of these conditions at the individual level. We were indeed able to identify markers that provide information about sleep quality (SE, TST, WASO, etc; Fig. 4) beyond binary classifications, potentially capturing the severity of objective markers of insomnia and offering a PSG inference of sleep quality. Moreover, even when our classifier made errors, we observed that these errors occurred for patients near the diagnosis cut-offs, indicating that individuals who are close

to these hard boundaries are logically more often misclassified (Fig. 4). Based on the new insights, we propose to move towards a continuous and multi-dimensional model of insomnia for which SOSP could represent one of the dimensions. This continuous approach could bypass the complex and unresolved issue of a common definition of what constitutes SOSP⁴¹. Our hypnogram-based approach could be further leveraged to find markers specific of different dimensions of insomnia (SOSP, physiological arousal, psychological arousal, chronotype, etc). Our analytic pipeline being light and fully automated, it is both accessible and scalable and could be routinely deployed in sleep clinics. In the future, it could also be deployed to hypnograms derived from other recordings than PSG recordings⁵⁴. Finally, a continuous and multi-modal approach enriched by features extracted from PSG recordings could help implement a precision medicine for insomnia⁵⁵.

Our analysis, though utilizing high-level metrics such as entropy and Kullback-Leibler divergence, could be effectively mapped to traditional and well-established EEG spectral features, and also to sleep metrics as SE, WASO and TST. This mapping not only validates our approach but also enhances its interpretability. For example, low-frequency EEG activity characterizes intrusions during wakefulness (i.e., high wake entropy), while high-frequency activity marks intrusions during sleep (high sleep entropy). Also, increases in high-frequency activity between consecutive epochs during deep sleep correlated with higher instability (high Kullback-Leibler divergence). These findings resonate with previous findings on local modulations of sleep and wakefulness^{20,26}. Additionally, our intrusions and instability features relate to dynamical macrostructural PSG features like self-transitions between wakefulness (Fig. 5D), which are also disrupted in insomnia^{31,56}. Unlike previous work that used black box deep learning models^{38,40}, our algorithm is fairly simple and can be explainable. This transparency facilitates clinical application, as it makes our findings easier for clinicians to interpret and utilize, and bridges the gap between advanced computational metrics and traditional sleep markers. Potential optimization could provide an estimate of the susceptibility of a given patient to psychological interventions as cognitive behavioral therapy for insomnia⁵⁷.

Previous findings on sleep and SOSP could explain the increase in entropy and instability that we observed in the two insomnia groups (Fig. 2). Indeed, an increase in faster oscillations is associated with the feeling of being awake while asleep (FAWS), as assessed directly when awakening individuals from sleep²³. In insomnia patients, this association between FAWS and faster oscillations is present in both NREM and REM sleep²³. It is likely that such increases in faster oscillations within sleep epochs would lead to an increase in the entropy of hypnograms similar to the one we report here. Such modulations of faster oscillations within sleep have also been associated with parasomnia³⁸, dreaming²⁸, and cognitive processes during sleep⁵⁹, which suggests links between local modulations of sleep and a broad range of sleep disruptions including insomnia with SOSP²⁶. Local modulations of cortical dynamics could thus help closing the gap between the objective and subjective assessment of sleep by stressing the importance of global and local dynamics within sleep^{20,24}. Our approach could help quantify these local changes beyond the standard global classification of sleep stages. These local dynamics could themselves be caused by a misalignment of sleep with the circadian phase, metabolic changes (e.g., higher metabolic activity during sleep in certain brain regions), or a disruption of the neuromodulation of sleep/wake transitions (e.g., a higher noradrenergic activity during sleep)²⁰. These factors have been shown to impact sleep^{60–63} and provide interesting avenues of research to better understand and treat insomnia^{10,51}.

While our study provides significant insights into insomnia, several limitations should be acknowledged. First, this is a mono-centric retrospective study although we controlled for sex, age, size, and BMI but groups were not matched for these variables. We could not account however for all comorbidities, which are common in insomnia⁶⁴. Nonetheless, this dataset is close to real clinical data recorded in sleep clinics, which eases its generalisability when translating these results into clinical applications. Additionally, our study lacked controlled wake recordings, preventing us from

assessing the level of sleep intrusions during normal daytime wakefulness outside of the period just before and during sleep. This limitation could explain our poor prediction of SOL, as we lacked information about pre-sleep activity. However, the strong predictive power and interpretability in terms of traditional EEG features suggest that our procedure is actually capturing distinct underlying physiological processes. In fact, when using an alternative automated sleep scoring algorithm (U-sleep, not trained with insomnia data), we found a bad performance, stressing the fact that our dedicated specific algorithm captures physiological relevant features of insomnia and sleep quality. Despite these limitations, our findings offer intriguing avenues for future research, particularly in exploring both hyperarousal and hypoarousal more comprehensively. Another limitation is that the differences we observe between SOSP– and SOSP+ relate to a single night of polysomnography per subject, which is compared to a subjective complaint of insomnia that covers a previous and longer period of time. This is in line with the data that can be available to practitioners (a subjective and chronic complaint of insomnia associated with a punctual assessment of sleep through a PSG). Our stance also corresponds to the classical definition of SOSP in clinical practice (e.g., the definition of “paradoxical insomnia”⁶⁵). While this approach is well-suited for clinical practice and the identification of biomarkers of SOSP, it is limited in its ability to uncover the underlying neural mechanisms of SOSP. To address the question of mechanisms more specifically, obtaining subjective assessments of sleep, particularly following spontaneous or induced awakenings, is preferable²³. In addition, our approach does not account for the night-to-night variability in PSG recordings, and phenomena such as the first-night or reverse first-night effects^{66,67}. It is thus possible that the SOSP+ and SOSP– labels could change from night to night at the individual level. However, our goal was to characterise the phenomenon of SOSP in general, not to draw detailed inferences on specific individuals. Future studies could employ other approaches to characterise SOSP for a specific night by comparing PSG recordings with subjective assessment of the same night. The same analytic approach as presented here could be applied.

Finally, we propose redefining the terminology used to SOSP in the scientific literature and in the information provided to patients. Terms such as “subjective insomnia”, “paradoxical insomnia” and SSM do not reflect the fact that patients with or without SOSP both follow the standard criteria for the diagnosis of insomnia. These terms also imply that the discrepancy is due to a problem of incorrect sleep perception and not an incorrect sleep quantification¹⁰. Terms such as “intrusions”, “instability” and “mixing” of sleep states can offer a more accurate description of sleep, its complexities and disorders^{68–70}. By correlating the subjective complaint with physiological data, we aim to contribute in bridging the gap between subjective reports and objective measurements. We favor the idea that SOSP is actually a mismeasurement rather than a misperception¹⁰, and that is a core dimension of insomnia (and sleep in general) rather than a sub-type of insomnia⁵⁰. Our work illustrates how subjective experience is a fundamental dimension of sleep²⁰, and taking it into account enables a better explanation of the inter-individual variability observed in neural activity. This integrative approach could provide a more graded and comprehensive understanding of sleep and consciousness research in general, enhancing both clinical and research perspectives.

Methods

Participants

The cohort examined in this study is an extension of a previously published study³¹. The total number of available recordings was 1042. However, 115 recordings were discarded because of noisy data or because the hypnogram was predicted with a balanced accuracy lower than 0.4 (see below). These recordings were excluded from all analyses, leading to a total sample size of 927 recordings. These excluded files are not included in the reported sample sizes or results. The remaining 927 participants had an age distribution of 44.4 ± 14.0 years and were 66.3% of females (Table 1 and Supplementary Table 9). Each participant underwent a single night of PSG to assess their sleep patterns. The recruitment strategy differed between patients with

Table 1 | Participants characteristics

Variables		SOSD- (n = 624)	SOSD+ (n = 199)	GS (n = 104)
Anthropometric data	Sex (% female)	69.7 (n = 435)	74.4 (n = 148)	30.8 (n = 32)
	Age (years)	46.8 ± 14.1	40.8 ± 12.1	36.2 ± 11.9
	BMI (kg/m ²)	23.1 ± 3.86	23.0 ± 3.56	24.2 ± 2.61
Comorbidities	Smoking (%)	21.8	23.6	27.3
	Depression (%)	25.0	22.0	0.0
	Anxiety (%)	19.0	19.4	0.0
	Cardiovascular disease (%)	13.8	5.2	0.0
	Pulmonary disease (%)	5.2	5.2	3.6
	Gastrointestinal disease (%)	4.4	4.0	0.0
	Infectious disease (%)	1.3	0.6	0.0
	Metabolic disease (%)	10.5	7.5	0.0
	Cancer (%)	3.7	1.8	0.0
	Covid (%)	3.3	3.2	0.0
	Epilepsy (%)	0.8	1.2	0.0
	Pain (%)	7.4	8.6	0.0
	Other comorbidity (%)	38.3	36.8	17.9
	Treatments	Antidepressant (%)	23.0	23.1
Antihistamine (%)		10.2	9.5	0.0
Antipsychotic (%)		3.3	4.1	0.0
Benzodiazepine (%)		25.0	24.9	0.0
Melatonin (%)		6.7	7.1	0.0
Z-drugs (%)		14.6	11.8	0.0
Opioid (%)		0.6	0.6	0.0
RLS medication (%)		1.2	0.6	0.0
Other medication (%)		41.1	39.6	3.6

chronic insomnia and good sleepers. PSG recordings from patients were obtained as part of their clinical care whereas PSG recordings from good sleepers were obtained as part of their participation in research protocols (see details below). In both cases, participants provided written consent and approved the re-use of their data for research purposes. Chronic insomnia was diagnosed following the International Classification of Sleep Disorders, Third Edition (ICSD-3)¹⁶, and defined as a complaint of persistent difficulties in sleep initiation, maintenance, or early morning awakenings for a minimum of three nights per week over a span of at least three months¹⁵. Patients were screened for obstructive sleep apnea (OSA) or periodic limb movement disorder (PLMD) by analysing PSG recordings and following existing guidelines (ICSD-3)¹⁶. All individuals with OSA and PLS were excluded from our analyses.

Insomnia without SOSD (SOSD -). The insomnia without SOSD cohort comprised 624 patients (mean age 46.8 ± 14.1 years and 435 females (~70%)), who were retrospectively identified from the Sleep and Vigilance Center at Hôtel-Dieu Hospital in Paris, France. As part of routine clinical evaluation, these patients underwent PSG (either laboratory-based or ambulatory) to assess their sleep. We leveraged PSG recording to identify patients with significant objective disruption of their sleep as measured by a single night of PSG: sleep onset insomnia (SOL > 30 min), OR sleep maintenance insomnia (WASO > 30 min), OR early morning awakening insomnia (natural awakening > 1 h before desired wake time)^{15,16}. Subjects meeting at least one of these criteria were categorized as “SOSD-.”

Insomnia with SOSD (SOSD+). 199 patients (mean age 40.8 ± 12.1 years and 148 females (~75%)) constituted the insomnia with SOSD group, identified through the same retrospective process at Hôtel-Dieu’s

Sleep and Vigilance Center. SOSD denotes a disjunction between perceived and objective sleep patterns. This discrepancy can be established in various ways, for example by comparing PSG metrics and subjective reports on the same night. Here, we compared a general complaint of poor sleep (falling within the diagnostic criteria of Chronic Insomnia) and the PSG results obtained in one night of sleep recording. Patients were included in the SOSD+ group if they exhibited relatively normal PSG metrics: SOL < 30 min, AND WASO < 30 min, AND natural awakening < 1 h before desired wake time. Subjects meeting all these criteria were categorized as “SOSD+ “ so that all insomnia patients were either SOSD- or SOSD+. This approach does not take into account the variability in PSG outcomes across nights but is well adapted to clinical settings where, typically, only one night of PSG is performed.

Control Group (GS). 104 participants (mean age 36.2 ± 11.9 and 32 females (~31%)) comprised the control group, recruited from prior VIFASOM studies^{71–80}, characterized by absence of subjective insomnia complaints, corroborated by PSG findings. The initial baseline night from prior protocols was consistently selected for analysis, conducted by the VIFASOM team. Control subjects underwent rigorous screening to exclude confounding factors such as sex, age, and comorbidities. Exclusions included individuals with OSA, PLMD, or TST outside the range of 4–8 h. In addition, subjects with chronic unbalanced pathologies, psychotropic treatments or those that could affect sleep were excluded from the protocols. Shift and night workers, people exposed to jet-lag or irregular sleep schedules (based on a 2-week sleep diary) were also excluded.

Clinical characteristics. In addition to undergoing PSG assessments, both SOSD- and SOSD+ subjects participated in comprehensive clinical

evaluations, which included a thorough medical history examination conducted by a qualified physician. Similarly, control subjects underwent a similar vetting process to ensure their suitability for study inclusion.

Comprehensive clinical data were systematically collected for all participants, encompassing various demographic variables such as sex and age, as well as anthropometric measures including height, weight, and body mass index. Additionally, detailed information regarding the participants' insomnia diagnosis, encompassing sleep initiation, maintenance, early morning awakening, and sleep state misperception, was recorded. The duration and nature of complaints, smoking status, current medication usage, and comorbidities were also documented (see Table 1 and Supplementary Table 1). Medications were categorized based on their known impact on sleep, while comorbidities were classified according to established associations with insomnia¹⁵. Disorders with fewer than 5 instances were grouped under the "other" category.

PSG recordings and sleep scoring

The study employed multiple PSG devices, including NOX-A1 (Nox Medical, $n = 820$), ACTIWAVE (CamNTEch Ltd, $n = 83$), MORPHEUS (Micromed S.p.A., $n = 13$), and SOMNOLOGICA (Medcare, $n = 11$). The NOX-A1, MORPHEUS, and SOMNOLOGICA systems included at least two EEG electrodes from C3, C4, F3, F4, O1, and O2, all referenced to their contralateral mastoid. The ACTIWAVE system used at least two of the EEG electrodes O1, O2, C3, F3, and FP1, also referenced to their contralateral mastoid. The number of EEG, EOG, and EMG derivations varied across devices, while respiratory parameters were recorded using NOX devices exclusively. All devices utilized in the study were validated and routinely employed in the sleep center and/or IRBA laboratory settings.

Following the completion of the night recording, PSG data underwent visual examination and scoring in line with the AASM guidelines³¹ to generate individual hypnograms. The identification of arousals, leg movements, and respiratory events was also conducted. Arousals were characterized by EEG acceleration within epochs lasting between 3 and 15 s. This visual scoring process was executed by a sleep medicine specialist. Specifically, 30-second epochs containing EEG (mastoid-referenced), EMG, and EOG data were categorized into wakefulness (W), NREM stages 1 to 3 (N1, N2, N3), and REM sleep. As part of this retrospective analysis, all PSG files were subsequently re-scored by the same experienced sleep technician, ensuring consistency in sleep scoring across the dataset. From this secondary scoring, key parameters reflecting sleep macro-structure were extracted from individual hypnograms, including TST, WASO, SOL and duration of each sleep stage. For a detailed description of the sleep macro-structure see Supplementary Tables 5, 6, 7 and 8.

Ethics

The research adhered to the ethical guidelines outlined in the Declaration of Helsinki of 1975, revised in 2001. Approval for the retrospective analysis of PSG data from both patients and controls was obtained from the local ethics committee (Comité de Protection des personnes (CPP) Ouest IV, Nantes, France). Patient data was pseudonymized and analyzed in compliance with legal regulations set forth by the Commission Informatique et Liberté (CNIL) in France. All ethical regulations relevant to human research participants were followed.

Data analysis

Preprocessing. EEG electrodes were referenced to its contralateral mastoid and then data was bandpass filtered in the 0.5–40 Hz range with a zero-phase finite impulse response with a Hamming window. Then, data was epoched in 30 s non-overlapping windows such that it matched the epochs used for the hypnogram. Only epochs with a valid sleep stage (W, N1, N2, N3, REM) were considered in the following analysis.

Spectral analysis. The power spectral density (PSD) was computed for each EEG electrode using the multitaper method implemented in the *mne* Python library⁸². Given that different EEG devices had different sampling

rates, all the PSDs were projected into a common frequency range between 0.5 and 40 Hz with intervals of 0.5 Hz. PSDs were computed only for epochs that were scored as W, N1, N2, N3 or REM. Noisy epochs were usually not included because they were not scored as one of these 5 states if the signal quality did not allow for accurate scoring. We did not perform further data cleaning. Then, for each valid epoch, we extracted the normalized the power in 6 canonical bands (delta: 0.5–4 Hz, theta: 4–8 Hz, alpha: 8–12 Hz, sigma: 12–16 Hz, beta: 16–30 Hz and gamma: 30–40 Hz). Normalization was performed by dividing the PSD values by the sum of the whole PSD.

Data harmonization by hypnodelensity estimation. The data used in this work was recorded with devices that differed in the number of electrodes, the electrode layout and the sampling rate. This diversity hinders the possibility of directly analyzing all the subjects in a common space. To overcome this limitation, and to use a sample as big as possible, we developed an harmonization procedure based on estimating hypnodeldensities (HD)³⁸. A HD is a probability distribution supported on the 5 different stages and quantifies the probability of each sleep stage to be present in a given epoch. This approach conveys more information than traditional hypnograms and provides a continuous representation of sleep dynamics.

To estimate the HD we developed an analytical pipeline based on massive feature extraction and machine learning. We used the *catch22* Python library⁴² to characterize each available EEG and EOG channel with 22 features (see the original publication for a detailed description). These features have been tailored to efficiently characterize a wide diversity time-series by capturing complementary and non-redundant aspects of the signals. Then, for example, if a recording had 3 EEG + 1 EOG channels, we would obtain $22 \times 4 = 88$ features. This analysis was performed for each epoch and the number of features changed for each subject depending on the number of EEG and EOG channels.

Next, we used an intra-subject leave-one-out classification approach to predict the sleep stage of each epoch using these *catch22* features and the respective hypnograms as input for a XGBoost multiclass classification algorithm⁸³ (see below for algorithm parameter details). For example, if a subject had 1000 epochs, we used the features and respective sleep stages of 999 epochs to train the classifier and tested on the remaining epoch. This procedure was repeated for each epoch. Finally, to obtain the HD, the classifier's prediction scores were transformed into class probabilities by the Platt scaling calibration procedure⁸⁴. This way we exploited and integrated the information present both in the PSG and in the hypnograms (scored by experts), enabling for the projection of every epoch of each subject into the same 5-dimensional space.

Intrusions and instability. As HD are probability distributions, tools from information theory can be leveraged to characterize them in a principled and interpretable way. Here we wanted to characterize two aspects of sleep dynamics: i) intrusions of other stages and ii) instability. We quantified them using two well known measures, the entropy (H) and the Kullback-Leibler divergence (D_{KL})⁸⁵, respectively. Entropy quantifies the uncertainty or surprise associated with the measurement of a given variable, while the D_{KL} measures the difference between two probability distributions. When base 2 logarithms are used, their units are maybits. In detail, for a random variable x with discrete probability distribution $P(x)$, the entropy $H(x)$ is defined as:

$$H(x) = - \sum_{i \in \Omega} p_i \log_2(p_i), \quad (1)$$

where x corresponds to the HD of a given epoch, $\Omega = \{W, N1, N2, N3, REM\}$, i.e., the 5 possible sleep stages, and p_i is the probability of each possible stage. For discrete probability distributions, entropy is a non-negative measure that is maximized when all the stages have the same probability (uniform distribution) and is 0 when a given stage has the maximum probability (a Dirac distribution). Then, an HD with large

intrusions of other stages will have high entropy (as the HD approaches a uniform distribution), while an epoch with small intrusions will have small entropy (as the HD approaches a Dirac distribution). This way, entropy is directly proportional to the level of intrusions and inversely proportional to the ‘purity’ of a given HD. In other words, the more certain we are about a specific stage, the lower the entropy.

Then, to measure the instability of an epoch with respect to the next one within the same sleep stage, we used the D_{KL} to quantify the difference between the two consecutive HD. In detail, for two probability distributions $P(x)$ and $Q(x)$ (i.e., the HD of two consecutive epochs) the D_{KL} follows:

$$D_{KL}(P||Q) = \sum_{i \in \Omega} p_i \log_2 \left(\frac{p_i}{q_i} \right), \quad (2)$$

where p_i and q_i is the probability of the stage i for the P and Q HD, respectively. If the HD of two consecutive epochs are equal, i.e., maximal stability, the D_{KL} is zero, while it increases as the two HD become different. Then, the D_{KL} is directly proportional to the instability of two consecutive epochs. In other words, if a big transition of the dynamics occurs from one epoch to the next one, the D_{KL} will be big.

Finally, for each subject, the average and standard deviation of both metrics were computed across all the epochs associated with a specific stage. This yielded a set of 16 features (the entropy and the D_{KL} for 4 stages and their respective average and standard deviation; N1 was not considered as a feature). These features were subsequently used as inputs to train a classifier for discriminating between GS, SODS– and SODS+, and also for a regression to predict TST, WASO, SOL and SE.

Feature selection. To enhance interpretability and diminish noise, we conducted a feature selection process on the set of information theoretic features. For both classification and regression tasks, we employed the minimum redundancy maximal relevance (MRMR) algorithm⁸⁶ to rank the features. This algorithm assesses feature relevance using the performance of a random forest, while redundancy is measured by the mean correlation with the rest of the features. Subsequently, the features were progressively included as inputs for cross-validated classification or regression based on their MRMR scores. To determine the final set of features, we computed the cross-validated balanced accuracy (for classification) or root mean squared error (RMSE, for regression) for each feature subset. We then identified the optimal subset as the point where further inclusion failed to significantly enhance performance (measured via a Wilcoxon rank sums test).

Machine learning classifier and regression. All the classification and regression tasks were performed using a XGBoost algorithm⁸³. XGBoost is a type of boosted ensemble of decision trees that have been proven accurate and efficient for many real-world applications. For multiclass problems (stage prediction and three-class prediction of diagnose) we used the following parameters: a *softprob* objective function, a *logloss* evaluation metric, a learning rate of 0.1, maximum tree depth of 6, a subsampling of 80% of subject, a feature subsampling of 80% and a regularization gamma parameter of 0.1. For the case of binary classifications (SODS– vs SODS+), the parameters were the same, with the only difference that the maximum tree depth was 3. For the regression tasks the objective function was the MSE and the rest of the parameters were the same as for classifications. These parameters were selected to avoid overfitting and, as they yielded good results, no hyperparameter optimization was done. For classification and regression, performance metrics were obtained from a repeated (60 iterations, unless other is specified) 5-fold stratified cross-validation procedure by generating 5 random partitions with approximately the same number of samples (subjects for diagnose prediction, epochs for sleep stage prediction). Then, the model was fit using 4 of these partitions, and its performance (balanced accuracy and confusion matrix for classification, and RMSE for

regression) was tested using the group that remained, and repeated this process until all partitions were used for testing and training. For the stage prediction (see *Data harmonization by hypnodensity estimation*), a leave-one-out procedure was used, so balanced accuracies and confusion matrices were computed after predicting each epoch separately.

Matching samples by size, sex, age and body-mass index. To verify that imbalances in size, sex, age and body-mass index (BMI) could be biasing the results (see Supplementary Table 1), we performed a subsample analysis where these variables were controlled. In the case of size, we used the smallest sample size among classes (GS = 104 for the three-class tasks and SODS+ = 199 for the SODS+ vs SODS– task) and took 30 equal sized random subsamples of the other class(es). This way the algorithm was trained with a perfectly balanced sample in terms of size.

For the binary classification between SODS– and SODS+, we matched the samples in terms of sex, age and BMI. To do so, we randomly selected a subsample with perfect balance in terms of sex and size (38 females, so 76 subjects for each condition) and used a Kolmogorov-Smirnov test to evaluate whether SODS– and SODS+ significantly differed both in age and BMI. If both tests were rejected (p -value > 0.1), the subsample was considered as balanced in size, sex, age and BMI. We generated 500 of these balanced subsamples and used them for the binary classification task.

Interpretation of information theoretic measures. To interpret the entropy and D_{KL} in terms of traditional sleep features (PSD and stage transitions) we performed correlations between entropy and the EEG spectral power and between the D_{KL} and self-transitions of sleep stages, respectively. For entropy and D_{KL} , we took the power in the 6 canonical bands (delta: 0.5–4 Hz, theta: 4–8 Hz, alpha: 8–12 Hz, sigma: 12–16 Hz, beta: 16–30 Hz and gamma: 30–40 Hz) of all the epochs related to a specific stage and correlated to the entropy and D_{KL} values associated with those epochs. We obtained one correlation value for each stage and subject. For the D_{KL} , we computed the transition matrix for each subject and extracted the self-transition associated with each stage and correlated with the average D_{KL} of that stage, obtaining one correlation value for the whole sample and for each condition.

Statistics and reproducibility

A Wilcoxon rank sums non-parametric test was used both to evaluate the differences between the information theoretic values and to compare the distributions of performance values obtained from the classifiers. Pearson correlation was used to compute correlation between entropy and D_{KL} and the macro and microstructural features of PSG. All p -values were corrected for multiple comparisons with the Bonferroni method. Group sample sizes correspond to SODS–: $n = 624$, SODS+: $n = 199$ and GS: $n = 104$; Each sample corresponds to a different participant. No replicates were used.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Due to protection of personal privacy, the clinical database used in this work cannot be publicly available. However, the data can be provided by DL pending scientific review and a completed material transfer agreement. Requests for the PSG recordings and associate metadata should be submitted to damien.leger@aphp.fr.

Code availability

Computational codes to extract hypnodensities and compute intrusions and instability from an example dataset are available in a Zeonodo repository⁸⁷.

Received: 27 August 2024; Accepted: 20 February 2025;
Published online: 12 March 2025

References

- Cirelli, C. & Tononi, G. Is sleep essential? *PLoS Biol.* **6**, e216 (2008).
- Ohayon, M. M. Epidemiology of insomnia: what we know and what we still need to learn. *Sleep Med. Rev.* **6**, 97–111 (2002).
- Taylor, D. J., Lichstein, K. L., Durrence, H. H., Reidel, B. W. & Bush, A. J. Epidemiology of Insomnia, Depression, and Anxiety. *Sleep* **28**, 1457–1464 (2005).
- Javaheri, S. & Redline, S. Insomnia and Risk of Cardiovascular Disease. *Chest* **152**, 435–444 (2017).
- Roth, T. Insomnia: Definition, Prevalence, Etiology, and Consequences. *J. Clin. Sleep Med.* **3**, S7–10 (2007).
- Léger, D. et al. Insomnia and accidents: cross-sectional study (EQUINOX) on sleep-related home, work and car accidents in 5293 subjects with insomnia from 10 countries. *J. Sleep Res.* **23**, 143–152 (2014).
- Léger, D. & Bayon, V. Societal costs of insomnia. *Sleep Med. Rev.* **14**, 379–389 (2010).
- Léger, D. & Poursain, B. An international survey of insomnia: under-recognition and under-treatment of a polysymptomatic condition. *Curr. Med. Res. Opin.* **21**, 1785–1792 (2005).
- Bastien, C. H. et al. Insomnia and sleep misperception. *Pathol. Biol.* **62**, 241–251 (2014).
- Stephan, A. M. & Siclari, F. Reconsidering sleep perception in insomnia: from misperception to mismeasurement. *J. Sleep Res.* **32**, e14028 (2023).
- Rundo, J. V. & Downey, R. Polysomnography. in *Handbook of Clinical Neurology* vol. 160 381–392 (Elsevier, 2019).
- Marino, M. et al. Measuring Sleep: Accuracy, Sensitivity, and Specificity of Wrist Actigraphy Compared to Polysomnography. *Sleep* **36**, 1747–1755 (2013).
- Carskadon, M. A. & Dement, W. C. Normal Human Sleep: An Overview. in *Principles and Practice of Sleep Medicine* 13–23 (Elsevier, 2005). <https://doi.org/10.1016/B0-72-160797-7/50009-4>.
- Riemann, D. et al. European guideline for the diagnosis and treatment of insomnia. *J. Sleep Res.* **26**, 675–700 (2017).
- Riemann, D. et al. The European Insomnia Guideline: An update on the diagnosis and treatment of insomnia 2023. *J. Sleep Res.* **32**, e14035 (2023).
- Sateia, M. J. International classification of sleep disorders-third edition: highlights and modifications. *Chest* **146**, 1387–1394 (2014).
- Center for Drug Evaluation and Research. Guidance for Industry: Guidelines for the clinical evaluation of Hypnotic Drugs. (1997).
- Guideline on medicinal products for the treatment of insomnia. (2011).
- Andrillon, T. Sleep: Feeling awake while asleep. *Curr. Biol.* **31**, R1578–R1580 (2021).
- Andrillon, T. & Oudiette, D. What is sleep exactly? Global and local modulations of sleep oscillations all around the clock. *Neurosci. Biobehav. Rev.* **155**, 105465 (2023).
- Valko, P. O., Hunziker, S., Graf, K., Werth, E. & Baumann, C. R. Sleep-wake misperception. A comprehensive analysis of a large sleep lab cohort. *Sleep Med* **88**, 96–103 (2021).
- Schinkelshoek, M. S., de Wit, K., Bruggink, V., Fronczek, R. & Lammers, G. J. Daytime sleep state misperception in a tertiary sleep centre population. *Sleep Med* **69**, 78–84 (2020).
- Stephan, A. M., Lecci, S., Cataldi, J. & Siclari, F. Conscious experiences and high-density EEG patterns predicting subjective sleep depth. *Curr. Biol.* **31**, 5487–5500.e3 (2021).
- Andrillon, T. How we sleep: From brain states to processes. *Rev. Neurol. (Paris)* **179**, 649–657 (2023).
- Kaplan, K. A. et al. When a gold standard isn't so golden: Lack of prediction of subjective sleep quality from sleep polysomnography. *Biol. Psychol.* **123**, 37–46 (2017).
- Siclari, F. & Tononi, G. Local aspects of sleep and wakefulness. *Curr. Opin. Neurobiol.* **44**, 222–227 (2017).
- Andrillon, T. & Oudiette, D. Global and Local Sleep. in *Advances in the Psychobiology of Sleep and Circadian Rhythms* 1–16 (Routledge, London, 2023). <https://doi.org/10.4324/9781003296966-1>.
- Siclari, F. et al. The neural correlates of dreaming. *Nat. Neurosci.* **20**, 872–878 (2017).
- Andrillon, T., Poulsen, A. T., Hansen, L. K., Leger, D. & Kouider, S. Neural Markers of Responsiveness to the Environment in Human Sleep. *J. Neurosci.* **36**, 6583–6596 (2016).
- Riedner, B. A. et al. Regional Patterns of Elevated Alpha and High-Frequency Electroencephalographic Activity during Nonrapid Eye Movement. *Sleep Chronic Insomnia: A Pilot Study.* *Sleep* **39**, 801–812 (2016).
- Andrillon, T. et al. Revisiting the value of polysomnographic data in insomnia: more than meets the eye. *Sleep Med* **66**, 184–200 (2020).
- Decat, N. et al. *Beyond Traditional Visual Sleep Scoring: Massive Feature Extraction and Unsupervised Clustering of Sleep Time Series.* <http://biorxiv.org/lookup/doi/10.1101/2021.09.08.458981> (2021) <https://doi.org/10.1101/2021.09.08.458981>.
- Perlis, M. L., Merica, H., Smith, M. T. & Giles, D. E. Beta EEG activity and insomnia. *Sleep Med. Rev.* **5**, 365–376 (2001).
- Fasiello, E. et al. Decreased Delta/Beta ratio index as the sleep state-independent electrophysiological signature of sleep state misperception in Insomnia disorder: A focus on the sleep onset and the whole night. *NeuroImage* **298**, 120782 (2024).
- Fulcher, B. D. & Jones, N. S. hctsa: A Computational Framework for Automated Time-Series Phenotyping Using Massive Feature Extraction. *Cell Syst.* **5**, 527–531.e3 (2017).
- Fiorillo, L. et al. Automated sleep scoring: A review of the latest approaches. *Sleep Med. Rev.* **48**, 101204 (2019).
- Rosenberg, R. S. & Van Hout, S. The American Academy of Sleep Medicine Inter-scoring Reliability Program: Sleep Stage Scoring. *J. Clin. Sleep Med.* **09**, 81–87 (2013).
- Stephansen, J. B. et al. Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. *Nat. Commun.* **9**, 5229 (2018).
- Vallat, R. & Walker, M. P. An open-source, high-performance tool for automated sleep staging. *eLife* **10**, e70092 (2021).
- Perslev, M. et al. U-Sleep: resilient high-frequency sleep staging. *Npj Digit. Med.* **4**, 1–12 (2021).
- Castelnovo, A. et al. The paradox of paradoxical insomnia: A theoretical review towards a unifying evidence-based definition. *Sleep Med. Rev.* **44**, 70–82 (2019).
- Lubba, C. H. et al. catch22: CAnonical Time-series CHaracteristics. *Data Min. Knowl. Discov.* **33**, 1821–1852 (2019).
- Dijk, D.-J. Slow-wave sleep deficiency and enhancement: Implications for insomnia and its management. *World J. Biol. Psychiatry* **11**, 22–28 (2010).
- Antelmi, E. et al. From state dissociation to status dissociatus. *Sleep Med. Rev.* **28**, 5–17 (2016).
- Castelnovo, A., Lopez, R., Proserpio, P., Nobili, L. & Dauvilliers, Y. NREM sleep parasomnias as disorders of sleep-state dissociation. *Nat. Rev. Neurol.* **14**, 470–481 (2018).
- Guthrie, R. S., Ciliberti, D., Mankin, E. A. & Poe, G. R. Recurrent Hippocampo-neocortical sleep-state divergence in humans. *Proc. Natl. Acad. Sci.* **119**, e2123427119 (2022).
- Emrick, J. J., Gross, B. A., Riley, B. T. & Poe, G. R. Different Simultaneous. *Sleep States Hippocampus Neocortex.* *Sleep* **39**, 2201–2209 (2016).
- Nir, Y. et al. Regional slow waves and spindles in human sleep. *Neuron* **70**, 153–169 (2011).
- Vilela, M. et al. Identifying time-resolved features of nocturnal sleep characteristics of narcolepsy using machine learning. *J. Sleep Res.* e14216. <https://doi.org/10.1111/jsr.14216> (2024).

50. Edinger, J. D. & Krystal, A. D. Subtyping primary insomnia: is sleep state misperception a distinct clinical entity? *Sleep Med. Rev.* **7**, 203–214 (2003).
51. Dressle, R. J. & Riemann, D. Hyperarousal in insomnia disorder: Current evidence and potential mechanisms. *J. Sleep Res.* **32**, e13928 (2023).
52. Bonnet, M. H. & Arand, D. L. Hyperarousal and insomnia: State of the science. *Sleep Med. Rev.* **14**, 9–15 (2010).
53. Riemann, D. et al. The hyperarousal model of insomnia: a review of the concept and its evidence. *Sleep Med. Rev.* **14**, 19–31 (2010).
54. Thomas, R. J., Wood, C. & Bianchi, M. T. Cardiopulmonary coupling spectrogram as an ambulatory clinical biomarker of sleep stability and quality in health, sleep apnea, and insomnia. *Sleep* **41**, zsx196 (2018).
55. Lim, D. C. et al. Reinventing polysomnography in the age of precision medicine. *Sleep Med. Rev.* **52**, 101313 (2020).
56. Wei, Y. et al. Sleep Stage Transition Dynamics Reveal Specific Stage 2 Vulnerability in Insomnia. *Sleep* **40**, zsx117 (2017).
57. Mitchell, L. J., Bisdounis, L., Balleisio, A., Omlin, X. & Kyle, S. D. The impact of cognitive behavioural therapy for insomnia on objective sleep parameters: A meta-analysis and systematic review. *Sleep Med. Rev.* **47**, 90–102 (2019).
58. Cataldi, J., Stephan, A. M., Marchi, N. A., Haba-Rubio, J. & Siclari, F. Abnormal timing of slow wave synchronization processes in non-rapid eye movement sleep parasomnias. *Sleep* **45**, zsac111 (2022).
59. Türker, B. et al. Behavioral and brain responses to verbal stimuli reveal transient periods of cognitive integration of the external world during sleep. *Nat. Neurosci.* **26**, 1981–1993 (2023).
60. Kay, D. B. et al. Subjective–Objective Sleep Discrepancy Is Associated With Alterations in Regional Glucose Metabolism in Patients With Insomnia and Good Sleeper Controls. *Sleep* **40**, zsx155 (2017).
61. Khazaie, H. et al. Functional reorganization in obstructive sleep apnoea and insomnia: A systematic review of the resting-state fMRI. *Neurosci. Biobehav. Rev.* **77**, 219–231 (2017).
62. Grimaldi, D. et al. Autonomic dysregulation and sleep homeostasis in insomnia. *Sleep* **44**, zsa274 (2021).
63. Lack, L. C., Micic, G. & Lovato, N. Circadian aspects in the aetiology and pathophysiology of insomnia. *J. Sleep Res.* **32**, e13976 (2023).
64. Sarsour, K., Morin, C. M., Foley, K., Kalsekar, A. & Walsh, J. K. Association of insomnia severity and comorbid medical and psychiatric disorders in a health plan-based sample: Insomnia severity and comorbidities. *Sleep Med.* **11**, 69–74 (2010).
65. Diagnostic classification of sleep and arousal disorders. 1979 first edition. Association of Sleep Disorders Centers and the Association for the Psychophysiological Study of Sleep. *Sleep* **2**, 1–154 (1979).
66. Newell, J., Mairesse, O., Verbanck, P. & Neu, D. Is a one-night stay in the lab really enough to conclude? First-night effect and night-to-night variability in polysomnographic recordings among different clinical population samples. *Psychiatry Res.* **200**, 795–801 (2012).
67. Riedel, B. W., Winfield, C. F. & Lichstein, K. L. First night effect and reverse first night effect in older adults with primary insomnia: does anxiety play a role? *Sleep Med.* **2**, 125–133 (2001).
68. Chouvarda, I. et al. Cyclic alternating patterns in normal sleep and insomnia: structure and content differences. *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.* **20**, 642–652 (2012).
69. Wassing, R. et al. Restless REM Sleep Impedes Overnight Amygdala Adaptation. *Curr. Biol. CB* **29**, 2351–2358.e4 (2019).
70. Riemann, D., Dressle, R. J., Benz, F., Palagini, L. & Feige, B. The Psychoneurobiology of Insomnia: Hyperarousal and REM Sleep Instability. *Clin. Transl. Neurosci.* **7**, 30 (2023).
71. Drogou, C. et al. Effects of Acute Caffeine Intake on Insulin-Like Growth Factor-1 Responses to Total Sleep Deprivation: Interactions with COMT Polymorphism - A Randomized, Crossover Study. *Lifestyle Genomics* **16**, 113–123 (2023).
72. Gomez-Merino, D. et al. Strategies to Limit Cognitive Impairments under Sleep Restriction: Relationship to Stress Biomarkers. *Brain Sci* **12**, 229 (2022).
73. Rabat, A. et al. Limited Benefit of Sleep Extension on Cognitive Deficits During Total Sleep Deprivation: Illustration With Two Executive Processes. *Front. Neurosci.* **13**, 591 (2019).
74. Bougard, C. et al. Motorcycling performance and sleepiness during an extended ride on a dynamic simulator: relationship with stress biomarkers. *Physiol. Meas.* **41**, 104004 (2020).
75. Faraut, B. et al. Daytime Exposure to Blue-Enriched Light Counters the Effects of Sleep Restriction on Cortisol, Testosterone, Alpha-Amylase and Executive Processes. *Front. Neurosci.* **13**, 1366 (2019).
76. Sauvet, F. et al. Beneficial effects of exercise training on cognitive performances during total sleep deprivation in healthy subjects. *Sleep Med* **65**, 26–35 (2020).
77. Arnal, P. J. et al. Sleep Extension before Sleep Loss: Effects on Performance and Neuromuscular Function. *Med. Sci. Sports Exerc.* **48**, 1595–1603 (2016).
78. Debellemaniere, E. et al. Performance of an Ambulatory Dry-EEG Device for Auditory Closed-Loop Stimulation of Sleep Slow Oscillations in the Home Environment. *Front. Hum. Neurosci.* **12**, 88 (2018).
79. Quiquempoix, M. et al. Effects of Caffeine Intake on Cognitive Performance Related to Total Sleep Deprivation and Time on Task: A Randomized Cross-Over Double-Blind Study. *Nat. Sci. Sleep* **14**, 457–473 (2022).
80. Erblang, M. et al. Genetics and Cognitive Vulnerability to Sleep Deprivation in Healthy Subjects: Interaction of ADORA2A, TNF- α and COMT Polymorphisms. *Life* **11**, 1110 (2021).
81. Berry, R. B. et al. Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J. Clin. Sleep Med. JCSM. Publ. Am. Acad. Sleep Med.* **8**, 597–619 (2012).
82. Gramfort, A. et al. MNE software for processing MEG and EEG data. *NeuroImage* **86**, 446–460 (2014).
83. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (2016)
84. Böken, B. On the appropriateness of Platt scaling in classifier calibration. *Inf. Syst.* **95**, 101641 (2021).
85. Entropy, Relative Entropy, and Mutual Information. in *Elements of Information Theory* 13–55 (John Wiley & Sons, Ltd, 2005). <https://doi.org/10.1002/047174882X.ch2>.
86. Ding, C. & Peng, H. Minimum redundancy feature selection from microarray gene expression data. in *Computational Systems Bioinformatics. CSB2003. Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003* 523–528 (IEEE Comput. Soc, Stanford, CA, USA, 2003). <https://doi.org/10.1109/CSB.2003.1227396>.
87. Herzog, R. Jupyter Notebook for Article ‘Sleep and wake intrusions: A continuous approach to explain insomnia and sleep state misperception’. <https://doi.org/10.5281/zenodo.13143110> (2024).

Acknowledgements

This survey has been supported by an unrestricted grant of the AID (Agence Innovation Défense) to the RAPID PANDORE-IA project and received financial support from Agence Nationale de la Recherche (ANR-22-CE37-0006-01) and the European Research Council (ERC-StG SleepingAwake 101116748).

Author contributions

Conceptualization: R.H., D.L., T.A.; Methodology: R.H., F.C., A.A., U.H., M.C., D.L., T.A.; Investigation: R.H., F.C., A.A., T.A.; Visualization: R.H.;

Supervision: D.L., T.A.; Writing-original draft: R.H., F.C., A.A., U.H., D.L., T.A.; Writing-review & editing: R.H., F.C., A.A., U.H., D.L., M.C., T.A.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-025-07794-6>.

Correspondence and requests for materials should be addressed to Thomas Andrillon.

Peer review information *Communications Biology* thanks Bernd Feige, Célyne Bastien and Elisabetta Fasiello for their contribution to the peer review of this work. Primary Handling Editor: Jasmine Pan. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025