npj | parkinson's disease

Article

Published in partnership with the Parkinson's Foundation

https://doi.org/10.1038/s41531-025-01105-w

Sleep stage mixing is associated with poor prognosis in early Parkinson's disease

Check for updates

Pauline Dodet[1,2] ✉, Emmanuel During[3], Isabelle Arnulf[1,2], Claudia Trenkwalder[4,5], Brit Mollenhauer[4,6], Friederike Sixel-Döring[4], Emmanuel Roze[2,7,8], Marie Vidailhet[2,7,8], Thomas Andrillon[2], Stéphane Lehericy[2,9], Jean-Christophe Corvol[2,7,8], Graziella Mangone[2], Smaranda Leu-Semenescu[1,2], Emmanuel Mignot[10], Poul Jennum[11] & Andreas Brink-Kjaer[12] ✉

This study assessed the prognostic value of probabilistic sleep staging in early PD by quantifying sleep stage uncertainty and mixing. Data from two longitudinal cohorts (DeNoPa and ICEBERG) included 280 patients with early PD (mean duration 21.1 months) and 158 matched controls. Sleep stages were scored automatically using U-Sleep, which provides for each 30-s epoch a probabilistic distribution of wakefulness (W), NREM stage N1, N2, N3, and REM sleep. Uncertainty (percentage of sleep stages with less than 80% certainty) and mixed stages were quantified. PD patients showed higher sleep stage uncertainty and wake–N3 mixing, especially those with REM sleep behavior disorder. Greater W–N3 mixing was associated with worse motor and cognitive scores at baseline and with faster motor decline over time. EEG features of this mixing suggested unstable deep sleep. These findings support the value of automated sleep analysis to detect microstructural sleep disruptions linked to PD progression.

Emerging evidence suggests that Parkinson's disease (PD) is a heterogeneous disorder resulting from a multifactorial pathogenic process[1]. Clinically, this heterogeneity is reflected in a progression of motor and non-motor symptoms that varies between individuals, leading to various degrees of disability and evolving over variable time periods. For example, PD associated with REM sleep behavior disorder (RBD) may represent a different phenotype than PD without RBD, indicating a more severe and widespread synucleinopathy[2]. One of the major challenges in this field today is to identify early biomarkers that predict individual risk for negative cognitive and motor outcomes, as this could have strong implications for future strategies in developing personalized treatment.

The study of sleep architecture using polysomnography (PSG) has gained attention as an innovative and non-invasive tool to explore the dynamic functioning of neural networks during wakefulness and sleep[3]. Sleep studies provide simultaneous and continuous information on nocturnal physiological processes relevant to PD, including (i) brain activity via the EEG, (ii) muscle control via the electromyogram (EMG), and (iii) eye movement control via the electrooculogram (EOG). EEG spectral analysis and characterization of sleep patterns throughout the night are well-established markers that reflect the activity of specific neuronal populations in discrete brain structures[4–6]. Sleep architecture is altered in some studies early in PD[7], but does not show substantial differences when scored only by international rules[8] (leading to abnormal sleep patterns, sometimes unnoticed by patients themselves). The exploration of PSG-based markers for phenotyping or prognostication has thus far been limited in patients with PD. First, electrophysiological analyses have mostly been studied in small, retrospective, or monocentric series, not always including appropriate clinical evaluation[3]. Second, most have used classical sleep study summary measures[9] or common sleep features such as spindles[10,11] or slow waves[12,13]. In advanced PD, specific EEG patterns, or the absence of them, predict cognitive decline[10,14] or worsening motor symptoms[13]. The underlying pathogenesis may be the deposition of alpha-synuclein aggregates in neural

[1]Service des Pathologies du Sommeil et Centre de Référence National des Narcolepsies et Hypersomnies rares, Assistance Publique-Hôpitaux de Paris-Sorbonne (AP-HP-Sorbonne), Hôpital la Pitié-Salpêtrière, Paris, France. [2]Sorbonne University, Paris Brain Institute (ICM), Inserm CNRS, Paris, France. [3]Department of Neurology, Mount Sinai School of Medicine, New York, NY, USA. [4]Paracelsus-Elena Klinik, Kassel, Germany. [5]Department of Neurosurgery, University Medical Center Goettingen, Goettingen, Germany. [6]Department of Neurology, University Medical Center Goettingen, Goettingen, Germany. [7]Center of Excellence of Neurodegenerative Disease (CoEN), AP-HP, Pitié-Salpêtrière Hospital, Paris, France. [8]Assistance Publique Hôpitaux de Paris, Department of Neurology, Clinical Investigation Centre Neuroscience, Pitié-Salpêtrière Hospital, Paris, France. [9]Department of Neuroradiology, Assistance Publique Hôpitaux de Paris, Hôpital Pitié-Salpêtrière, Paris, France. [10]Department of Psychiatry and Behavioral Sciences, Center for Sleep Sciences and Medicine, Stanford University, Stanford, CA, USA. [11]Danish Center for Sleep Medicine, Glostrup University Hospital, Glostrup, Denmark. [12]Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark. ✉e-mail: pauline.dodet@aphp.fr; andbri@dtu.dk

circuits that regulate sleep and wakefulness[15]. Finally, the presence of RBD or RSWA was unknown in these studies, making it impossible to adjust the results for this parameter, which is one of the most important predictors of poor prognosis[10,13,14].

The scoring and analysis of PSG is complex and time-consuming, particularly in individuals with neurodegenerative diseases. However, recent studies have shown that several aspects of sleep analysis can be automated, including sleep staging[16,17], arousals detection[18], sleep spindles[19], and automated quantification of REM sleep without atonia (RSWA)[20]. Utilization of these methods could facilitate automated quantification of sleep-based markers of neurodegeneration. In addition, several methods based on deep learning have modeled RBD directly from EEG and EOG data[21,22]. Another innovative concept in sleep analysis is the use of hypnodensity, a new approach to analyzing sleep as a probabilistic distribution of sleep stages rather than distinct stages, providing a more fluid and dynamic view. Hypnodensity based on automatic scoring assesses the likelihood of a particular stage being present at any given time, which has been shown to follow the distribution of multiple experts on the same data[23]. This method complements traditional analysis, which assigns a single sleep stage to each 30-s period. It therefore provides a more nuanced representation of sleep, which has been shown to contain diagnostically relevant information for sleep apnea[17] and narcolepsy type 1[16]. This paves the way for the development and use of new metrics to quantitatively reflect abnormalities in the structure of sleep as a whole. This property is of particular interest in the context of neurodegenerative pathologies, and therefore PD, as a 'dissociated' aspect of sleep has often been reported. In this study, we choose U-sleep[22–25] as a probabilistic algorithm as it was extensively validated across 8 held-out clinical cohorts and is easily accessible, which allows our methods to be reproduced and applied for similar analyses. We defined two groups of metrics based on the hypnodensity graph, which described sleep stage uncertainty and sleep stage mixing. Sleep stage uncertainty was defined as the percentage of 30-s epochs in which the highest probability was less than 80%. Uncertainty was computed globally across all epochs (global uncertainty) and for epochs corresponding to a specific sleep stage (stage-dependent uncertainty) as determined by U-Sleep. Mixing between two stages was based on the product of the probabilities of these stages in each epoch, summed over the entire night.

Here, we aimed to (i) investigate destructuration of sleep in two large and distinct controlled cohorts of de novo and early PD with and without RBD, based on the sleep uncertainty and the mixing between sleep stages using hypnodensity derived from U-Sleep; (ii) and to evaluate how these metrics were associated with motor and cognitive scores at baseline and follow up.

Results

Demographic and clinical characteristics of the groups

A total of 280 participants with PD (126 in the DeNoPa cohort and 154 in the ICEBERG cohort) and 158 healthy controls (99 in the DeNoPa cohort and 59 in the ICEBERG cohort) were enrolled. Comparisons of demographic and clinical measures in the PD and control groups are shown in Table 1. In the whole group at baseline, the mean age at inclusion for patients with PD was 63.4 ± 9.6 years and 63.9 ± 8.0 years for controls. In the PD group, the mean disease duration (from the time of first symptom) was 18.1 ± 13.3 months in the ICEBERG cohort and 24.7 ± 33.1 months in the DeNoPa cohort. Patients were treated with a median total levodopa equivalent daily dose of none in the DeNopa cohort and 297.0 ± 269.4 mg in the ICEBERG cohort. Follow-ups of motor and cognitive tests were included over 4 years following the PSG test. Specifically, participants were followed for 1.1 ± 0.1, 2.1 ± 0.2, 3.1 ± 0.2, and 4.3 ± 0.3 years in the ICEBERG cohort and 2.2 ± 0.2 and 4.2 ± 0.2 years in the DeNoPa cohort. Participants did not necessarily complete all follow-up visits, a point which is detailed in Table 1. The same information is detailed separately for the DeNoPa cohort and the ICEBERG cohort in Supplementary Tables 1 and 2. Moreover, specific motor sub-scores of the MDS-UPDRS are displayed in Supplementary Table 3.

Comparison between U-sleep and manual scoring

The concordance of U-Sleep with human expert scoring was evaluated in both datasets separately for PD (Supplementary Table 4) and controls (Supplementary Table 5). On average, U-Sleep had an agreement with manual scoring in PD, yielding to an F1 score of 0.63 ± 0.10 in ICEBERG and 0.55 ± 0.10 in DeNoPA. It performed not better or worse in controls in each cohort, despite the fact that the U-sleep algorithms were not trained on PD data[17]. Importantly, the U-sleep algorithms do not use EMG to score PSG, removing an important confounder in patients with RBD. The largest discrepancies were found in the scoring of N1 sleep (F1 score < 0.22) in both cohorts and N3 sleep in the DeNoPa cohort (F1 score of 0.4 for both PD and controls).

The distribution of sleep characteristics computed by U-Sleep and manual scoring showed different interpretations of sleep (Supplementary Table 6). Most notably, the distribution of N1% and N3% were different across cohorts according to manual scoring, e.g., N1 in the PD group was 5.9 ± 5.1% of total sleep time in the ICEBERG cohort and 23.6 ± 11.0% in the DeNoPa cohort ($p < 0.0001$). Interestingly, between-cohorts discrepancies in N1 and N3 percentages were decreased and even reversed in some groups when using U-Sleep, e.g., N1 in PD was 8.9 ± 5.6% in the ICEBERG cohort and 7.9 ± 5.7% in the DeNoPa cohort ($p > 0.05$).

Sleep characteristics

The sleep characteristics (mainly based on U-Sleep scoring) are shown in Table 2, which highlights differences between the PD(-RBD) group, PD(+RBD) group, and control group. No effects of sleep macrostructure were significant in both cohorts, however, combined the PD(+RBD) group did present with lower REM sleep percentage overall. Moreover, the PD(+RBD) group presented significantly higher percentage of REM sleep without atonia than both the PD(−RBD) groups and controls.

Sleep stage uncertainty

The global uncertainty was lower in controls (33.2 ± 8.4%) than in the PD (-RBD) group (37.1 ± 11.0%) and the PD (+RBD) group (46.7 ± 13.5%). There was a gradual increase in all stage-dependent uncertainties from controls to PD(−RBD) and then to PD(+RBD) (Fig. 1 and Supplementary Table 7), except for N3 uncertainty, which did not show significance. The most pronounced increase in uncertainty was observed during REM sleep from 20.1 ± 14.4% in the control group to 28.4 ± 21.3% in the PD(-RBD) group and then to 52.6 ± 28.6% in the PD(+ RBD) group ($p < 0.0001$). There was a higher sleep stage-dependent uncertainty in N1, N2, and REM sleep in the PD(+RBD) group than in the control group across both cohorts.

Sleep stage mixing

Sleep stage mixing values are maximal for epochs containing high probabilities of two sleep stages occurring simultaneously. Across both cohorts, the global sleep stage mixing prevalence was higher for all combinations except for the combinations N1-N2 and N2-N3 in the PD(+RBD) group compared to the control group (Fig. 1 and Supplementary Table 7). Across both cohorts, the global mixing stage prevalence was higher for the W-N2, W-N3, and N2-REM combination in PD(-RBD) participants compared to controls.

Hypnodensity-based microsleep markers were correlated to examine how much they constitute separate measure. The correlation matrix is displayed in Supplementary Fig. 1, which shows moderate to strong correlations for the majority of measures.

Conventional metrics of sleep microstructure

The distributions of measures based on spectral analysis of EEG and sleep-wake events (arousals, spindles, and slow waves) are shown in Supplementary Table 8. The arousal index during NREM was lower in PD(+RBD) groups than in the control group. The arousal index during REM sleep was higher in PD(+RBD) the control groups. The sleep spindle density was lower both in PD(−RBD) and PD(+RBD) groups than in controls. The group comparisons of slow-to-fast ratios and slow wave measures were non-significant in at least one cohort.

Table 1 | Overview of demographics, medication, and clinical tests for controls and participants with Parkinson's disease in both the DeNoPa and Iceberg cohort combined

	n	Controls 158	PD 280	p-value
Demographical data				
Age [years]	μ ± σ	63.9 ± 8.0	63.4 ± 9.6	n.s.
Sex (female)	% (n)	40.6% (63)	36.1% (101)	n.s.
BMI [kg/m^2]	μ ± σ	26.2 ± 4.4	26.2 ± 4.3	n.s.
Parkinson's disease duration [months]	μ ± σ	-	21.1 ± 24.5	-
REM sleep behavior disorder	% (*n*)	0.0% (0)	25.2% (70)	<0.0001
Number of visits	μ ± σ	2.1 ± 1.0	2.2 ± 1.2	n.s.
Follow-up duration at last visit [years]	μ ± σ	3.6 ± 1.2	3.2 ± 1.5	0.037
Treatment				
Antidepressants	% (n)	4.6% (7)	11.4% (32)	0.03
Benzodiazepines	% (n)	2.0% (3)	1.4% (4)	n.s.
LEDD [mg]	μ ± σ	-	163.3 ± 248.5	-
Slope LEDD [mg per year]	μ ± σ	-	100.7 ± 95.6	-
Motor and cognitive symptoms				
MDS UPDRS III total score OFF (0–132)	μ ± σ	2.3 ± 3.8	26.2 ± 10.6	<0.0001
MDS UPDRS total score (0–199)	μ ± σ	6.3 ± 6.7	42.0 ± 15.7	<0.0001
Slope MDS UPDRS III total score [per year]	μ ± σ	0.26 ± 1.2	1.2 ± 3.7	0.0013
Slope MDS UPDRS total score [per year]	μ ± σ	0.9 ± 2.2	3.2 ± 5.6	<0.0001
MoCA total score (0–30)	μ ± σ	26.7 ± 2.6	26.5 ± 2.8	n.s.
Slope MoCA total score [per year]	μ ± σ	−0.086 ± 0.69	−0.32 ± 1	0.027

MDS-UPDRS movement disorder society-sponsored revision of unified Parkinson's disease rating scale, *MoCA* Montreal cognitive assessment. An overview for each individual cohort is displayed in Supplementary Tables S2 and S3.

Table 2 | Baseline comparisons of sleep recordings between PD with and without REM sleep behavior disorder and controls groups based on U-sleep scoring

	n	Controls 158	PD(-RBD) 210	PD(+RBD) 70	*p* value Control vs. PD(-RBD)	Control vs. PD(+RBD)	PD(-RBD) vs. PD(+RBD)
Sleep measures (U-Sleep)							
Total sleep time [min]	μ ± σ	356.1 ± 61.3	364.0 ± 63.2	355.9 ± 61.1	n.s.	n.s.	n.s.
Sleep efficiency	μ ± σ	78.6 ± 12.3	77.4 ± 11.8	74.9 ± 11.0	n.s.	0.0064	n.s.
WASO [min]	μ ± σ	71.5 ± 46.7	80.2 ± 53.0	86.0 ± 45.4	n.s.	0.0061	n.s.
Sleep latency [min]	μ ± σ	20.2 ± 17.3	17.7 ± 15.9	22.5 ± 19.5	n.s.	n.s.	n.s.
Sleep stages (U-Sleep, % of total sleep time)							
N1	μ ± σ	10.0 ± 6.2	8.4 ± 5.7	8.4 ± 5.6	0.0042	0.02	n.s.
N2	μ ± σ	54.2 ± 8.7	57.2 ± 11.1	58.9 ± 11.6	0.022	0.0067	n.s.
N3	μ ± σ	16.1 ± 7.9	16.9 ± 9.8	17.6 ± 10.5	n.s.	n.s.	n.s.
REM	μ ± σ	19.8 ± 7.3	17.4 ± 7.5	15.1 ± 8.4	0.0096	<0.0001	0.025
Manual scoring							
PLMI (n/h)	μ ± σ	26.2 ± 28.3	13.8 ± 22.0	30.7 ± 41.9	**<0.0001**	n.s.	0.0057
AHI (n/h)	μ ± σ	5.1 ± 7.9	5.1 ± 7.3	5.0 ± 7.8	n.s.	n.s.	n.s.
RSWAI tonic	μ ± σ	5.0 ± 7.1	10.0 ± 15.7	39.8 ± 25.7	0.0092	**<0.0001**	**<0.0001**

Comparisons between the 3 groups were evaluated using Kruskal-Wallis tests, and if significant, Mann–Whitney U tests were performed for pairwise comparisons. Comparisons were repeated within each cohort separately, and if both tests were significant ($p < 0.05$), the *p*-values are displayed in **bold**. Periodic leg movement index, Apnea-hypopnea index, and REM sleep without atonia index are based on manual scoring of the polysomnography.
WASO wake after sleep onset, *PLMI* periodic leg movement index, *AHI* apnea-hypopnea index, and *RSWAI REM* sleep without atonia index.

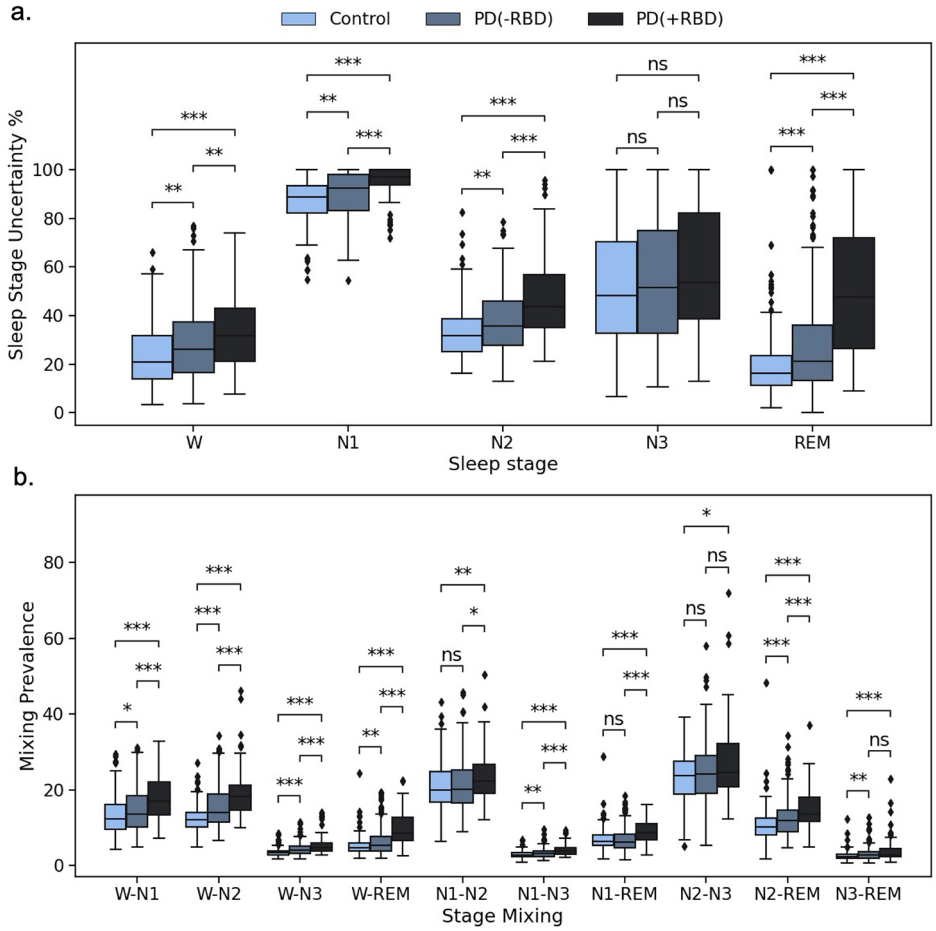

Fig. 1 | Comparison of hypnodensity-based sleep stage dependent uncertainty and mixing prevalence between PD participants with and without RBD and controls. Hypnodensity-based sleep stage dependent uncertainty (**a**) and mixing prevalence (**b**) in controls (light blue), PD(-RBD) participants (gray blue), and PD(+RBD) participants (black). Comparisons between the 3 groups were evaluated using Kruskal–Wallis tests, and if significant, Mann–Whitney U tests were performed for pairwise comparisons. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Hypnodensity measures and clinical symptoms in Parkinson's disease

The associations between altered sleep structure (derived from hypnodensities) and motor and cognitive symptoms in PD were assessed at baseline and follow-up, controlling for confounders (including demographics, medications, and RBD) (Table 3).

N1 uncertainty was the strongest predictor of MDS-UPDRS III score with an effect of 1.93 (0.79, 3.16) per standard deviation increment. For MDS-UPDRS total score, the amount of W-N3 mixing had the largest effect of 2.84 (1.04, 4.63) per standard deviation increment. Several measures were associated with lower MoCA, most notably N2 uncertainty as well as the amount of W-N3 and N1-N3 mixing.

To predict the evolution of motor and cognitive symptoms, the slope (a linear fit based on available follow-up tests) was modeled using each measure, while controlling for confounders and baseline test scores, thereby modeling additional effects in annual change. The amount of W-N3 and W-REM mixing was associated with worse motor outcomes (Table 3). The amount of W-N3 mixing predicted worse outcome (0.72 [0.01, 1.44] for MDS-UPDRS total score and 0.54 [0.09, 0.99] for MDS-UPDRS III total score per standard deviation increment). None of the effects were significant for predicting change in MoCA after controlling for confounders. The association between prevalence of W-N3 mixing and both baseline and annual change in symptom scores is shown in Fig. 2. A secondary analysis was performed for W-N3 mixing in subscores of the MDS-UPDRS (Supplementary Fig. 2), including the tremor score, PIGD score, axial score, and AR score. Here, W-N3 mixing was associated with worse AR score (1.13 [0.33, 1.93] per standard deviation increment); furthermore, it predicted worse annual change in the AR score (0.34 [0.04, 0.63] per standard deviation increment) and the axial score (0.06 [0.00, 0.12] per standard deviation increment).

Conventional metrics of sleep microstucture and clinical symptoms in Parkinson's disease

The associations between conventional sleep metrics (sleep stage percentages, slow-to-fast ratios, arousal index, spindle metrics, and slow wave metrics) and clinical motor and cognitive testing were also assessed (Supplementary Table 9). At baseline, no associations were found. The slow-to-fast ratio in wakefulness was associated with the slope of MDS-UPDRS total score (1.18 [0.53, 1.84] per standard deviation increment) and the slope of MDS-UPDRS III total score (0.67[0.25, 1.09] per standard deviation increment) but not with the slope of MoCA. Slow wave amplitude and sleep spindle frequency were significantly associated with cognitive score progression, with the strongest effect for spindle frequency of 0.24 [0.10, 0.37] per standard deviation increment.

Combined analysis of all sleep metrics

Linear lasso regression was used to determine which combination of features (both hypnodensity-based and conventional-based) best described motor and cognitive scores (Fig. 3). The strongest effects for baseline motor symptoms were the amount of W-N2 and W-N3 mixing as well as N1 uncertainty. For evolution of motor symptoms, the strongest combined effects included slow-to-fast ratio during wakefulness, W-N3 mixing, and W-REM mixing. For evolution of MoCA scores, the spindle frequency and slow wave amplitude were the strongest predictors.

Electrophysiological analysis of W-N3 mixing

In the analysis of EEG frequency content, we considered each 30-s epoch to belong to a single sleep stage, or mixed sleep stage if either the individual sleep stage probability was above 80% or if the frequency of sleep stage mixing was above 80%, respectively. In participants with PD, 466,213 W epochs, 208,540 N3 epochs, and 495 W-N3 epochs were analyzed. The

Table 3 | Associations between hypnodensity mixing features and clinical test scores at baseline and their yearly progression

Baseline	MDS UPDRS Total	MDS UPDRS III	MoCA Total
Sleep stage uncertainty			
Global	2.17 (0.23, 4.12)*	1.33 (0.03, 2.63)*	−0.28 (−0.61, 0.06)
W	2.06 (0.34, 3.77)*	1.28 (0.13, 2.43)*	−0.036 (−0.34, 0.27)
N1	2.78 (0.94, 4.62)**	1.93 (0.70, 3.16)**	−0.05 (−0.38, 0.28)
N2	1.96 (−0.01, 3.92)	1.05 (−0.27, 2.37)	−0.44 (−0.78, −0.09)*
REM	1.91 (0.02, 3.79)*	0.82 (−0.46, 2.10)	−0.26 (−0.60, 0.08)
Global sleep stage mixing prevalence			
W-N2	2.67 (0.76, 4.59)**	1.54 (0.25, 2.82)*	−0.24 (−0.57, 0.09)
W-N3	2.84 (1.04, 4.63)**	1.42 (0.20, 2.63)*	−0.38 (−0.71, −0.05)*
N1-N3	1.89 (0.06, 3.71)*	0.85 (−0.38, 2.08)	−0.40 (−0.72, −0.08)*
Slope [per year]			
Global sleep stage mixing prevalence			
W-N3	0.72 (0.01, 1.44)*	0.54 (0.09, 0.99)*	−0.0018 (−0.14, 0.14)
W-REM	0.72 (−0.03, 1.46)	0.67 (0.20, 1.13)**	−0.031 (−0.17, 0.11)

Effects per standard deviation in feature are calculated by linear regression, adjusting for age, sex, BMI, REM sleep behavior disorder, duration of PD symptoms, levodopa equivalent daily dose, antidepressant use, benzodiazepine use, and cohort. Yearly progression was also adjusted for baseline test score and yearly change in levodopa equivalent daily dose. Features with all non-significant association to clinical tests ($p > 0.05$) are not shown. $^{**}p < 0.01$, and $^{*}p < 0.05$.

W-N3 epochs above 80% were present in 22/280 subjects with PD and were transitioned to N3 (63.7%), W (33.3%), or N2 (3.0%). As shown in Fig. 4, W-N3 mixing on average follows N3 for delta-, theta-, alpha-, but was highest in beta activity compared to N3. Interestingly, beta activity within W-N3 mixing epochs showed larger temporal variance, indicating fluctuating high-frequency activity. There was no increase in the sigma activity in W-N3 compared to N3. As the W-REM mixing also showed some promise, we also included a spectral analysis of the mixed stages in supplementary Fig. 3. It shows that it broadly represents an average shift to higher frequencies while maintaining a similar within-epoch variance of frequencies as REM sleep.

Discussion

In this study, we used the automated sleep scoring by U-Sleep in two large and distinct early PD cohorts and used it to investigate changes in sleep structure based on EEG and EOG. Among participants with PD, especially those with RBD, we observed an increased prevalence of sleep stage uncertainty and mixing across all combinations of traditional sleep stages. Additionally, several salient markers of sleep stage uncertainty and mixing were associated with more severe motor and cognitive scores at baseline, as well as with their progression over time.

An innovative aspect is the use of automatic sleep scoring (U-Sleep) to analyze the global sleep structure and provide hypnodensity (a probabilistic distribution of sleep stages)[24], from which new features such as sleep uncertainty and sleep stage mixing have been calculated[16]. So-called "hypnodensity" approach sleep as a continuous, heterogenous, and dynamic process providing insights into a dissolution of the normal boundaries between sleep stages, even when other parameters seem to be preserved. One of the advantages of this concept is that it standardizes the way in which sleep recordings are scored, which is an undeniable asset in view of the significant inter-scorer variability in neurodegenerative diseases, a variability which, however, is strongly linked to the experience of scorers in this field[25]. As this is an algorithm trained on more traditional sleep data, it may fail to recognize the stage of sleep that a trained expert would have identified. But the interest in using this automatic scoring in our study lies not so much in the ultimate stage chosen by the algorithm as in its ability to assess the 'difficulty' of scoring this stage of sleep and to be able to extract metrics reflecting it.

This study revealed an extended and global alteration in sleep structure in early PD participants compared to controls, characterized by heightened uncertainty and increased sleep stage mixing. This phenomenon was observed both globally and across all sleep stages, including both NREM and REM sleep. These findings were consistent in both cohorts and with and without treatment. Although data on sleep architecture are scarce and inconsistent in early PD[26–30], previous studies have reported changes in macrostructure with increased sleep onset latency[26,29], decreased sleep efficiency and REM sleep stage, and increased N1 stage[26,31]. Changes in NREM microstructure have also been observed using spectral analysis, including a reduction in the low delta frequency range (0.78–1.2 Hz)[32], a reduced amount of slow wave activity[33], and concomitant increased scalp EEG alpha and sigma activities in NREM sleep[34] in small groups (<10 patients) of de novo PD patients compared to controls. In advanced PD, key figures that characterize NREM sleep, such as spindle density or K-complexes, are altered[35,36]. A recent study validated an extended set of scoring guidelines in neurodegenerative diseases, which included seven additional sleep-wake stages[37]. We speculate that the mixed sleep stages observed here may represent some of these additional stages, or could capture the heterogeneity of the standard sleep stages[38], thereby automatically providing a more accurate description of the complex sleep-wake physiology. In future studies, it may be interesting to compare the mixing of sleep stages with these new scoring guidelines.

We also investigated whether changes in sleep structure (based on hypnodensity and on conventional measures including spindle metrics and EEG frequencies) could reflect disease severity and could serve as electrophysiological markers of poor disease progression. Few studies have examined the association between electrophysiologic markers of sleep and clinical progression, but not in a large multicenter cohort of early patients with PD. First, we found a gradient of severity in sleep stage uncertainty and mixing, with PD participants experiencing RBD showing the most pronounced changes across all sleep stages. This was particularly evident in the mixing of REM-based sleep characteristics and was most pronounced in the mixing between REM sleep and wakefulness, which was expected in a REM sleep disorder. More interesting, this change in sleep structure was also observed in NREM sleep. Several studies have proposed that PD associated with RBD represents a distinct phenotype compared to PD without RBD, indicating a more severe and widespread synucleinopathy with broader brain damages[2,39,40]. Clinically, RBD in patients with PD is associated with poorer motor and cognitive performance[41], more severe nonmotor symptoms[41], and faster disease progression[42,43]. Our findings of increased sleep alterations are consistent with these findings. Second, we found that both NREM sleep uncertainty and NREM-Wake mixing were associated with worse baseline motor and cognitive scores and a more rapid progression of MDS-UPDRS III score. Most prominently, global W-N3 mixing prevalence had the largest effect. This feature predicted worse symptoms trajectories for MDS-UPDRS III total score. This is consistent with a retrospective study in 129 patients showing that the total amount of slow wave activity in N2 and N3 was related to worse motor progression in PD[13]. Regarding cognitive function, an association with sleep stage uncertainty, including N2 uncertainty, N1–N3 mixing, and W-N3 mixing prevalence, was only found at baseline. We found no associations between sleep stage uncertainty and mixing on the one hand, and the longitudinal slope of the MoCA total score on the other hand. Only higher spindle frequencies and slow wave amplitude were associated with better longitudinal progression of the MoCA score. Several studies have reported that electrophysiologic abnormalities in PD are predictive of cognitive decline. Impaired sleep spindle activity and slowing of posterior and temporal EEG frequencies during REM sleep and wakefulness are associated with cognitive worsening[14]. Schreiner et al. found that slow wave activity (1–4 Hz) was

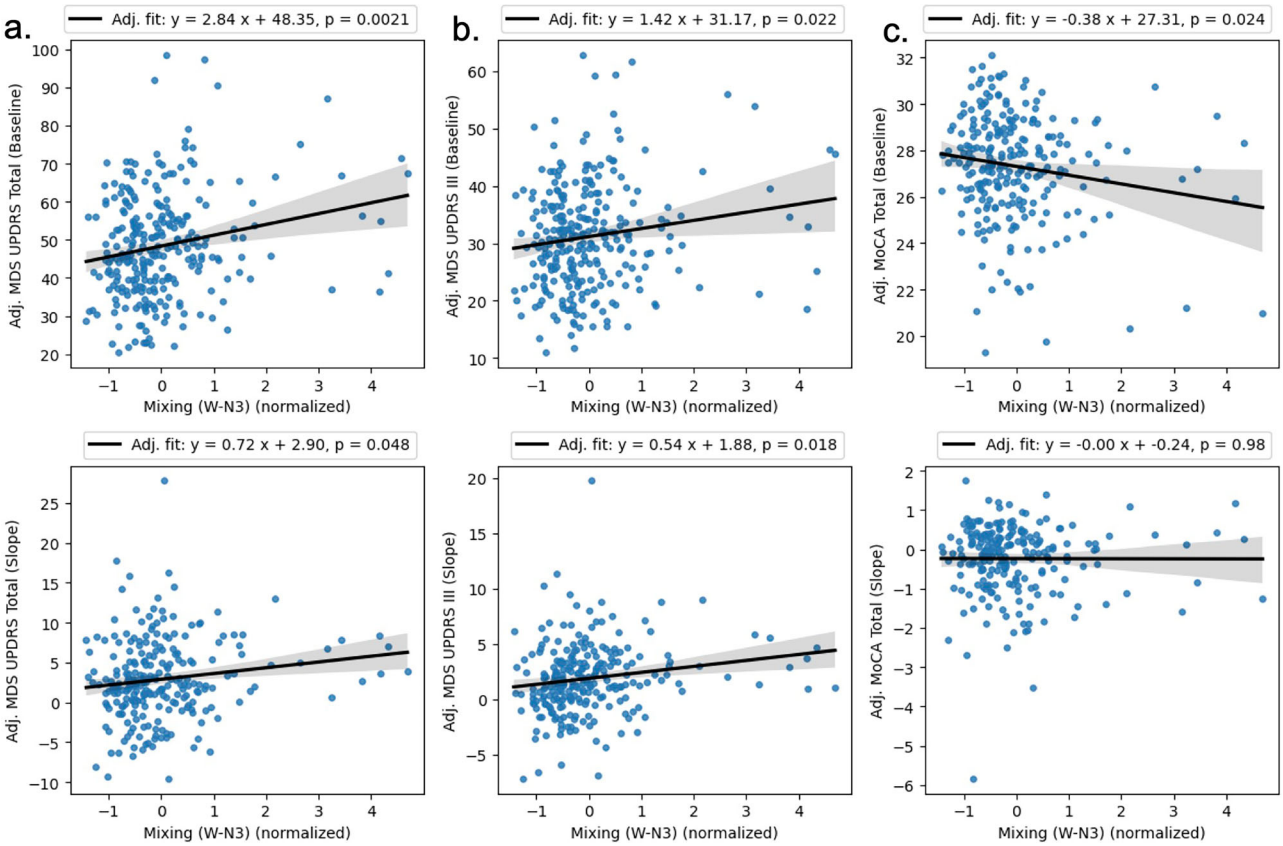

Fig. 2 | Relationship between the prevalence of W-N3 mixing and clinical symptoms in participants with Parkinson's disease. Partial residual plots showing the relationship between the prevalence of W-N3 mixing (normalized) and motor **a**, **b** and cognitive (**c**) symptoms (baseline and longitudinal slope) adjusted for confounders (described in Table 3). Regression lines in these plots correspond to those in Table 3, and the translucent bands indicate the 95% confidence interval of the fit. The partial residual was calculated including the intercept.

associated with higher MoCA scores in a cross-sectional study, and that this association was region dependent, with the strongest effect in frontal EEG derivations[12]. Nevertheless, a study of 68 patients with advanced PD showed that the slow waves characteristics were not associated with a poorer cognitive outcome[10]. The difference with our result could be explained by the different method of analysis and an earlier PD stage in our study[14]. Indeed, we used MoCA score as a continuous measure instead of a categorical measure (with and without dementia). Finally, the 4-year longitudinal follow-up in participants with high MoCA scores at baseline may have been insufficient to detect a change in MoCA scores.

Our final goal was to better characterize the physiological significance of the main marker that we highlighted. The global prevalence of W-N3 mixing likely represents instability in deep slow wave sleep, i.e., not a reduction of deep NREM sleep (as the percentage of N3 was not decreased in PD here) but rather that N3 sleep is disturbed and unstable. W-N3 mixing on average follows N3 sleep for low frequencies and is closer to beta activity for wakefulness. Moreover, beta activity within W-N3 mixing epochs shows larger temporal variance, indicating fluctuating high-frequency activity. Activity in the beta frequency band during wakefulness is associated with -among other things- voluntary motor drive. Its decrease is related to the execution of movements, and its increase is related to the cancellation of planned movements[44]. The excessive synchronization in the beta frequency band in the cerebral cortex and the cortico-basal ganglia nuclei in PD[45,46] has been proposed to be antikinetic in nature[47]. Indeed, the strongest clinical correlate of beta power in PD is the severity of bradykinesia[48]. In contrast, a significant decrease of beta activity has been observed in subthalamic local field potentials recorded through electrodes in N2 and N3 sleep, compared to wakefulness[49]. In our study, however, beta activity was increased in N3 sleep and was, for the first time, associated with a poorer motor and overall prognosis. This raises the question of the integrity of sleep functions associated with N3 sleep and how this abnormal N3 stage impacts the PD pathogenesis. Indeed, emerging evidence suggests that sleep could play a critical role in the pathophysiology of neurodegenerative diseases, through the elimination of toxic protein aggregates from the brain via the glymphatic system[50,51]. For example, in Alzheimer's disease, a strong bidirectional relationship between sleep disruption and increased amyloid-β deposition has been established[52–54]. An early work suggests that the efficacy of glymphatic fluid transport timely correlates with the amount of slow wave activity during sleep, which reflects the integrity of N2-N3, while flow stops at the onset of wakefulness[55,56]. In PD, few studies suggest that sleep alterations may be associated with faster progression[13,57], although the exact role of the glymphatic system during sleep on the clearance of aggregated α-synuclein remains unclear[58]. Our study provides additional evidence that impaired NREM sleep is associated with poorer motor prognosis. We speculate that this N3 dysfunction may impair function related to the glymphatic system. Further studies investigating this aspect by correlating with brain imaging are needed to evaluate this point. Alternatively, a common process may alter both N3 and the motor and cognitive networks.

This study with detailed demographic, clinical, and sleep measures provides a large, representative, systematic, and longitudinal follow-up of two distinct cohorts of participants with de novo and early PD over 4 years. To date, studies investigating sleep in PD and demonstrating discernible changes in different sleep stages have been limited by small sample sizes, significant heterogeneity in disease duration (abnormalities appearing gradually or increasing over time), lack of information on RBD status, on medications, and inconsistent analytical methods. A strong point here is that statistical analyses have been adjusted for confounders (age, sex, body mass index, disease duration, treatments, and the presence of RBD defined

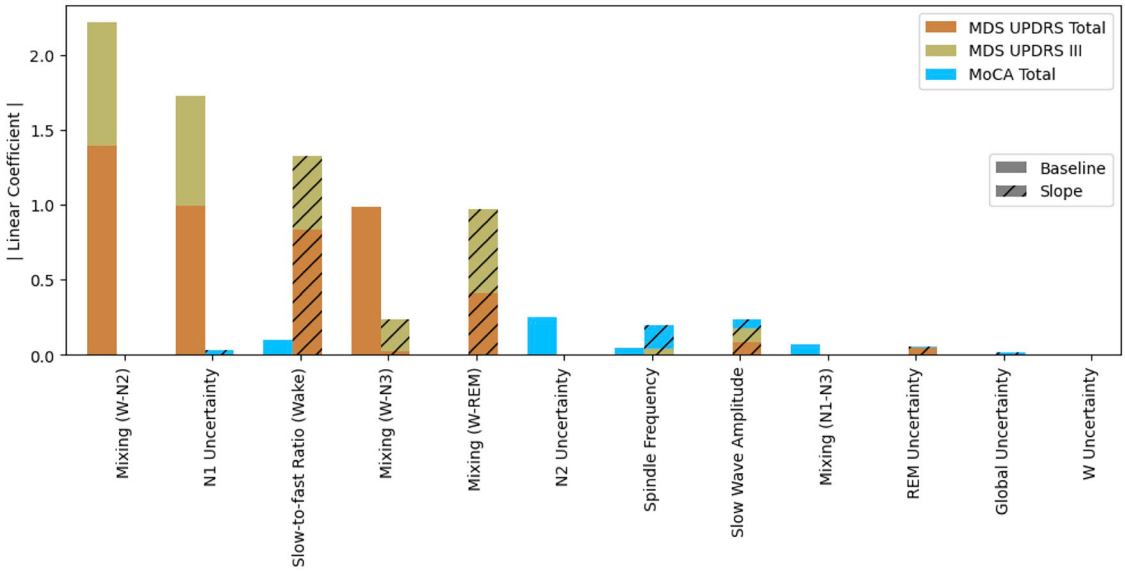

Fig. 3 | Ranking of measures based on the sum of the absolute value of their coefficients in modeling both baseline test scores and their longitudinal slope using a linear lasso regression. The model further included and adjusted for age, sex, BMI, REM sleep behavior disorder, duration of PD symptoms, levodopa equivalent daily dose, antidepressant use, benzodiazepine use, and cohort (yearly progression was also adjusted for baseline test score and change in levodopa equivalent daily dose).

by independent experts). We chose to adjust for RBD because it is well known to be associated with a poor outcome[43]. In addition, automatic sleep scoring offers objectivity and consistency in the scoring process, facilitating comparisons between data from different clinics and devices. Sleep scoring in patients with overt neurodegenerative diseases can be challenging using traditional guidelines, hence, some researchers have proposed alternative scoring criteria[37]. Consistent with this idea, we found significant differences in sleep stages between the two cohorts when scored manually by sleep experts. The change in distribution of N1 and N3 from manual to automatic scoring highlighted this homogenization, as the latter showed a much closer distribution. While the agreement was equal between PD participants and controls, the model uncertainty was higher in PD with and without RBD. To verify the robustness of our data and the additional contribution of automatic analysis, we also performed more conventional signal analysis methods using spectral analysis and sleep pattern characterization. Moreover, in a combined lasso model, we found that the sleep stage uncertainty and mixing provided complementary information to the previously studied conventional measures in modeling motor and cognitive symptoms at both baseline and their progression.

The main limitation of this study is that automatic analysis was based on a reduced subset of electrodes available in all PSGs, which reduced resolution and did not allow spatial analysis of microstructural changes. It might explain why F1 scores obtained here are lower than results reported in previous U-Sleep uses[16,17]. Future studies could also examine region/electrode-specific alterations, as other authors have suggested they could have predictive value[14]. Furthermore, neither the ECG nor the eye movement artifacts were eliminated. These artifacts are known to significantly impact power estimates in the delta band (0.5–4 Hz). Since delta power is a component of the slow-to-fast ratio used to quantify EEG slowing, these artifacts could bias the interpretation of EEG spectral slowing. Additionally, the objective of this study was not to validate the U-Sleep algorithm for scoring sleep recordings in patients with Parkinson's disease. To achieve this, a structured comparison with visual scoring from multiple experts would be required to compare U-Sleep's agreement with the inter-rater variability.

This research highlights the potential of automated scoring for objective and consistent assessment of sleep microstructure changes in this disorder, providing valuable insights into the intricate relationship between sleep and neurodegeneration. Further exploration of these alterations may provide important markers of disease progression and severity, and stimulate future investigations (and interventions) into the role of sleep in the pathophysiology of PD.

Methods

Participants and data collection

Participants were recruited from two independent longitudinal cohorts consisting of patients with de novo PD (DeNoPa cohort, Kassel, Germany) and early PD (ICEBERG cohort, Paris, France). Both studies performed a video polysomnography (PSG) at baseline and focused on examining the clinical course of the participants. Inclusion and exclusion criteria have been described previously in refs. 59,60. Briefly, in the DeNoPa cohort, patients were (1) newly diagnosed with PD according to UK Brain Bank criteria (at least two of resting tremor, rigidity, and bradykinesia); (2) aged between 40 and 85 years; (3) not exposed to levodopa during the 4 weeks prior to study enrollment and exposed for less than 2 weeks if previously exposed. In the ICEBERG cohort, patients had to be (1) between 18 and 75 years of age, (2) be diagnosed with PD according to the UK Brain Bank criteria, (3) have no or minimal cognitive impairment (defined as a Mini-Mental State Examination score greater than 26/30), (4) have a disease duration (from the time of first symptom) of less than 4 years. Controls were included based on the absence of current or past neurological or psychiatric disorders. There was no habituation night in the ICEBERG cohort and two nights in the KASSEL cohort, and we used data from the second night.

Standard protocol approvals, registrations, and patient consents

The DeNoPa cohort was registered at the German Registry for Clinical Trials (DRKS00000540) according to the WHO Trial Registration Data Set and the ICEBERG cohort at ClinicalTrials.gov (NCT02305147). The Ethics Committee at the Ârztekammer Hessen, Germany (approval no. FF89/2008) approved the DeNoPa study, and the Ethics Committee (CPP- Ile de France-Paris 6, RCB 2014-A00725-42) approved the ICEBERG cohort. All participants gave written informed consent to participate in the study.

Clinical assessment

All participants were examined by movement disorder specialists at baseline and with follow-up every two years in the DeNoPa cohort and annually in the ICEBERG cohort for 4 years. Patients participated to face-to-face interviews to record current medications (dopaminergic drugs use was converted to levodopa equivalent daily doses [LEDDs]) and clinical

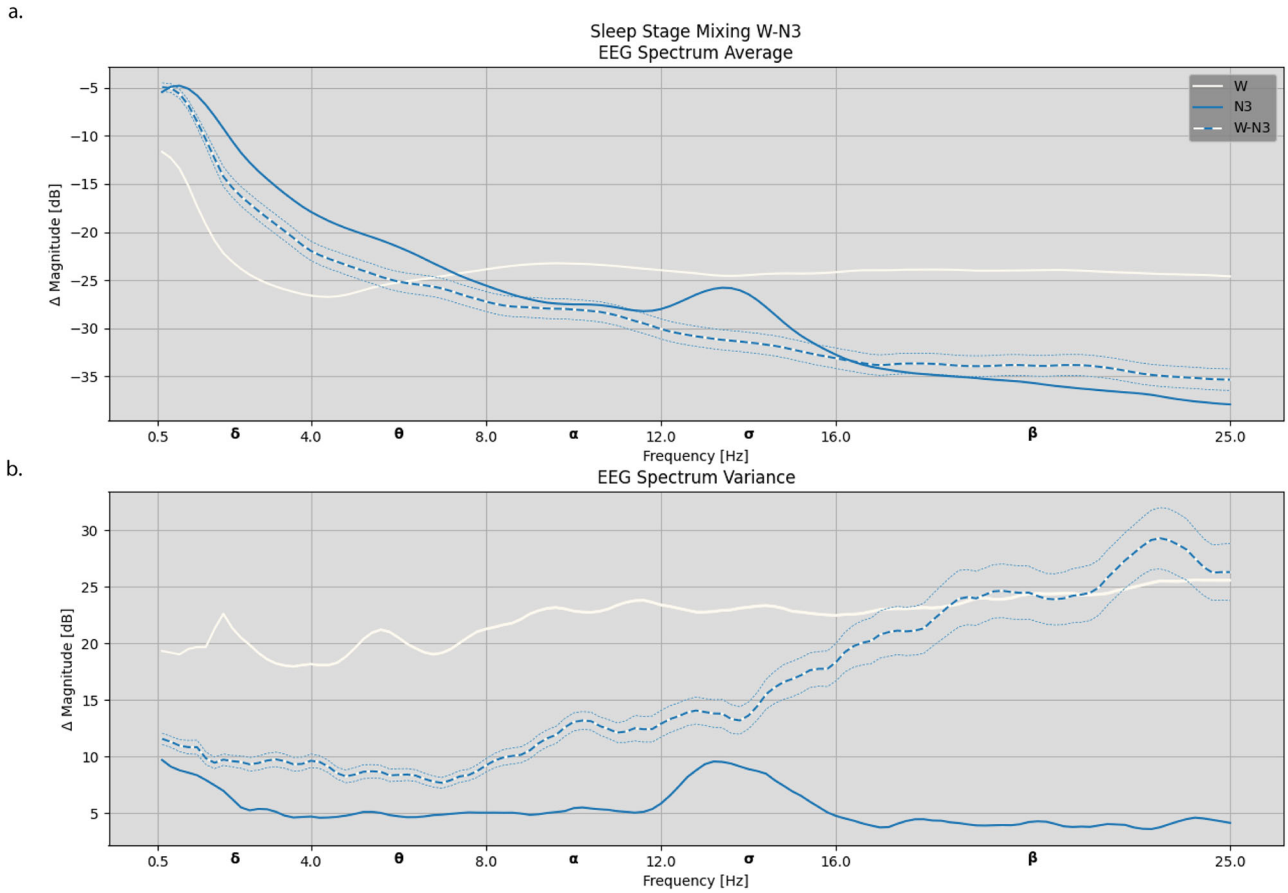

Fig. 4 | Distribution of EEG frequency content across mixed W-N3 sleep epochs determined by U-Sleep compared to Wake and N3 sleep in all participants with Parkinson's disease. The upper panel displays the average frequency power (**a**), and the lower panel shows the average frequency power variance **b** (i.e., how the frequency magnitude changes within 30-s epochs). The thinner lines indicate the mean 90% confidence interval.

measures included motor assessments using the Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS), and cognitive evaluations using the Montreal Cognitive Assessment (MoCA) and a sleep study every two years. Analyses of MDS-UPDRS subscores were performed following previously used definitions for the tremor score[61], postural instability/gait dysfunction (PIGD) score[62], axial score[63], and akinetic-rigid (AR) score[61].

Sleep measures

Patients underwent a sleep study while on usual medications in the sleep unit of Pitié-Salpêtrière Hospital, Paris, France, or of the Elena Klinik, Kassel, Germany. The recordings included 3 (F1, C3, O1, referenced to A2) electroencephalogram (EEG) channels, 2 electro-oculograms (EOG), chin and bilateral tibialis anterior muscles electromyograms (EMG), electro-cardiogram (EKG), oronasal flow (via oronasal thermistor plus nasal pressure), respiratory thoracic and abdominal efforts (via plethysmography), transcutaneous oxyhemoglobin, position, infrared video, and sounds. Sleep neurologists scored sleep stages, arousals, periodic leg movements (PLM), and respiratory events according to international criteria[8]. The presence of RBD was diagnosed through video-PSG by experienced raters and followed the international criteria.

Analyses of hybrid sleep stages

The version 2.0 of U-Sleep was used to extract the hypnodensity, which is a probabilistic distribution of sleep and wake stages for each[30]-second epoch of PSG recordings, including W, N1, N2, N3, and R[24]. The subset of electrodes available in all PSGs (C3-A2 EEG and left and right EOG) was used. An additional preprocessing step was added to remove power-line interference at 50 and 60 Hz using a second-order notch filter with a quality factor of 30, which was applied forward and backward.

Table 4 | Description of hypnodensity features. Here, H represents the hypnodensity of length N and S symbols the most certain sleep stages in H

Hypnodensity features	Formula
Global Sleep Stage Uncertainty Computed across all stages	$100 \cdot \frac{1}{N}\sum_{i=1}^{N}[\max(H_i)<0.9]$
Sleep Stage Dependent Uncertainty Computed for stage $\boldsymbol{s}_1$	$100 \cdot \frac{1}{\lvert S=s_1 \rvert}\sum_{i\in(S_i=s_1)}[H_{i,s}<0.9]$
Global Sleep Stage Mixing Prevalence Computed between stages $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$	$100 \cdot \frac{1}{N}\sum_{i=1}^{N} 2 \cdot \left(H_{i,s_1} \cdot H_{i,s_2}\right)^{\frac{1}{2}}$

The standard metrics of sleep architecture (including the total sleep time and the percentage of time spent in each sleep stage) were obtained from U-Sleep, as well as sleep stage mixing and uncertainty derived from hypnodensities. Each 30-s epoch is assigned a stage based on highest probability[16]. Sleep stage uncertainty was defined as the percentage of 30-s epochs in which the highest probability was less than 80%. Uncertainty was computed globally across all epochs (global uncertainty) and for epochs corresponding to a specific sleep stage (stage-dependent uncertainty) as determined by U-Sleep. Mixing between two stages was defined as the product of the probabilities of these stages in each epoch, summed over the entire night. The formula for each feature is shown in Table 4.

Examples of hypnodensities and derived features are shown in Fig. 5 for control, PD without RBD [PD-(RBD)], and PD with RBD [PD(+RBD)] participants. To ensure that these examples were representative, we selected

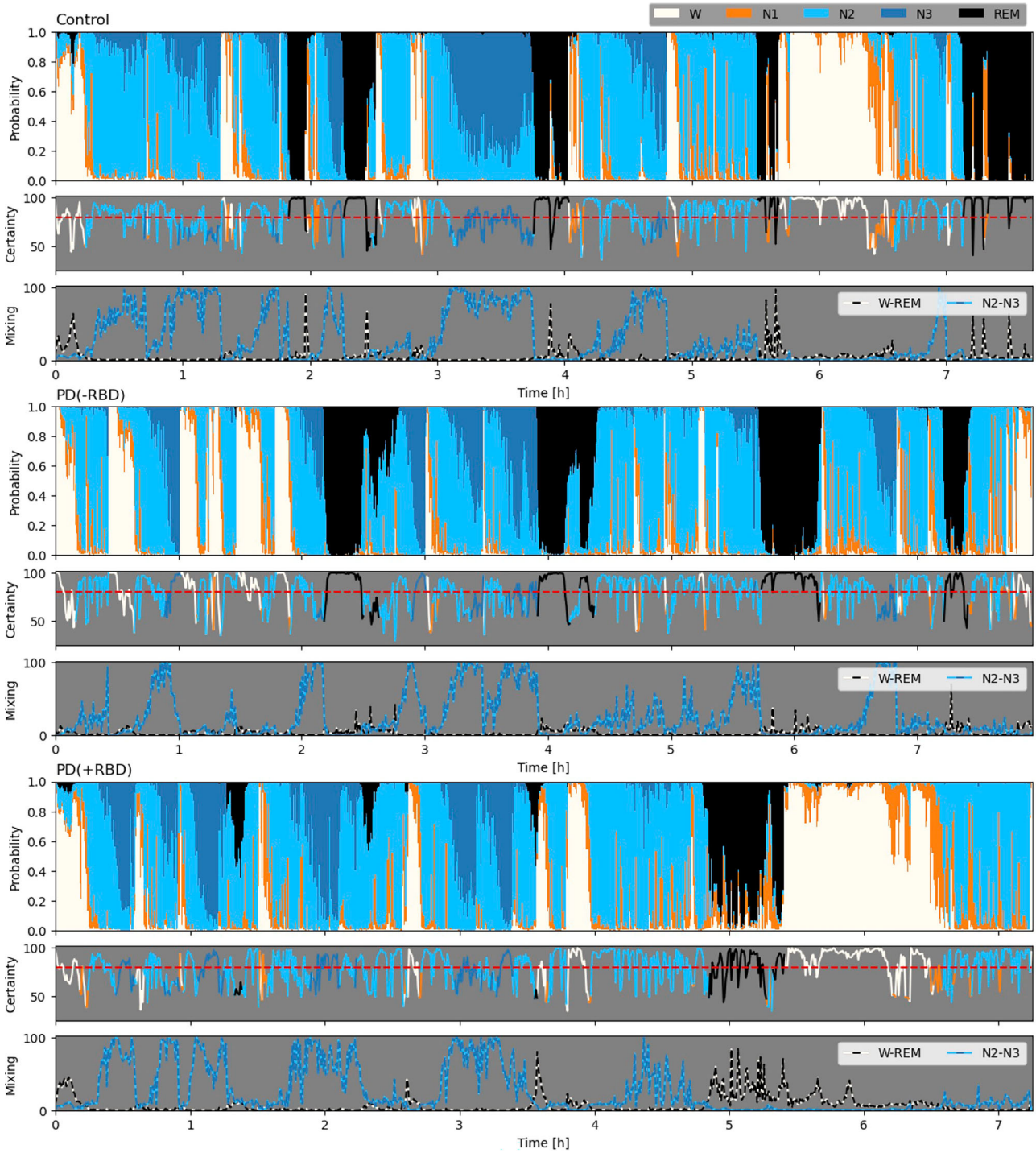

Fig. 5 | Hypnodensities modeling scored by U-Sleep in a control participant, as well as in a patient with Parkinson's Disease (PD) without rapid eye movement sleep behavior disorder (RBD) [PD(-RBD)], and a patient with PD and RBD [PD(+RBD], all being the closest to the median of their group. Representative examples of hypnodensities are shown with derived sleep stage uncertainty (with the red dotted line indicating the 80% threshold) and mixing for W-REM and N2–N3 for each subject. Note how mixed REM sleep is prominent in the PD + RBD subjects compared to the control subjects.

individuals who were closest to their group median across all features (weighed equally through z-score normalization) of sleep stage uncertainty and mixing as well as traditional sleep metrics (total sleep time, sleep efficiency, wake after sleep onset, sleep latency, N1%, N2%, N3%, REM %).

Analysis of conventional sleep metrics

A series of automatic methods was used to analyze the EEG spectral dynamics and detect cortical arousals, sleep spindles, and slow waves, in order to replicate previously found associations between sleep measures and clinical trajectories in early PD, and to investigate if hypnodensity-based measures provide complementary information to model motor and cognitive symptoms.

EEG frequency content was analyzed using the multitaper method[29] in 5-s windows without overlap and a c-parameter of 20. Slowing of EEG frequencies were estimated following Latreille et al.[13]. by computing a slow-to-fast ratios ([delta 0.5–4 Hz + theta 4–8 Hz]/[alpha 8–12 Hz + sigma 12–16 Hz + beta 16–30 Hz]) in wake, N2 plus N3 sleep, and REM sleep as scored by U-Sleep. Artifact removal was implemented for body movements

using the YASA (Yet Another Spindle Algorithm)[19] package. Here, 30-s epochs containing a 5-s window exceeding a three standard deviation threshold, which was calculated per sleep stage, were excluded from further analyses. Moreover, slow wave activity was extracted as the average relative power of delta activity across N2 and N3 sleep (excluding epochs with body movements) following Schreiner et al.[12]. Slow wave energy was defined as slow wave activity multiplied by hours of N2 + N3 sleep (counting all epochs).

Cortical arousals were scored using the Multimodel Arousal Detector[64], which processes central EEG, left and right EOG, and chin EMG. Multimodel Arousal Detector scores wakefulness and arousals with a 1-s resolution and combines them as a single measure. This improves the temporal resolution of longer arousals or wake periods slightly longer than 15 s, which traditionally are annotated as wakefulness if they are aligned with the 30-s scoring window.

Sleep spindles and slow waves were scored using YASA[19], which analyzes the relative power different frequencies, EEG morphology, and relative positions to score events. These events were labeled in NREM 2 and 3 sleep as scored by U-Sleep. Moreover, any sleep spindle detected during an automatically scored arousal was excluded. Spindle and slow wave activity were characterized as density (number/hours of N2 and N3 NREM sleep), amplitude (μV), and frequency (Hz, median frequency extracted through Hilbert transform).

Statistical analyses

Statistics were computed in Python v.3.7.12 using various software packages, including statsmodels v.0.12.2, scipy 1.7.3, and numpy 1.21.6. Group comparisons were performed using Mann–Whitney U and Kruskal–Wallis tests for continuous distributions. Chi-squared tests were used for binary variables. The level of significance was set at 0.05. The star notation indicates a significance level below 0.05 (*),0.01 (**), and 0.001(***). If the effect was significant, post hoc tests were performed separately in each cohort to determine whether the effect was present in both cohorts independently of each other.

The rate of longitudinal progression of motor and cognitive symptoms was assessed by calculating the slope of change in clinical scores per year for each clinical test. The slope was calculated with a fixed offset at baseline and using all available follow-up visits (DeNoPa: Year2 and Year4, ICEBERG: Year1, Year2, Year3, and Year4).

Comparisons between measures and clinical tests were performed using multiple regression to adjust for confounding variables, including age, sex, body mass index (BMI), medications, RBD, and cohort. The clinical test measure at baseline was adjusted for when testing for longitudinal progression of test score. The specific effect of W-N3 mixing prevalence was visualized using partial residual plots, where the partial residual was calculated as $\varepsilon_i + \beta_0 + \beta_x X$, where ε_i is the residual for observation i, β_0 is the intercept, and $\beta_x X$ is the effect of the variable of interest. A Lasso linear model was used to select a set of features that best related to each clinical test at baseline and for its longitudinal progression. These models initially included all the previously listed confounding variables as well as all for the features that had a significant association with either the baseline test score or its longitudinal slope. The Lasso model was optimally adjusted after 10-fold cross-validation. Examining the correlation of hypnodensity-based markers was assessed using Spearman's correlation. Spectral content of W, N3, and mixed W-N3 was examined using similar frequency analysis like that used to compute slow-to-fast ratios. Confidence intervals of mean frequency content were computed assuming a Gaussian distribution of frequency magnitudes.

Data availability

Polysomnography recordings and other clinical data from the ICEBERG and DeNoPa cohorts used in this study were subject to data sharing agreements but are available upon reasonable request from any qualified investigator.

Received: 30 July 2025; Accepted: 1 August 2025;
Published online: 26 September 2025

References

1. Berg, D. et al. Prodromal Parkinson disease subtypes - key to understanding heterogeneity. *Nat. Rev. Neurol.* **17**, 349–361 (2021).
2. Postuma, R. B. et al. REM sleep behavior disorder and neuropathology in Parkinson's disease. *Mov. Disord.* **30**, 1413–1417 (2015).
3. Zahed, H. et al. The neurophysiology of sleep in Parkinson's disease. *Mov. Disord.* **36**, 1526–1542 (2021).
4. Staresina, B. P. et al. Hierarchical nesting of slow oscillations, spindles and ripples in the human hippocampus during sleep. *Nat. Neurosci.* **18**, 1679–1686 (2015).
5. Fernandez, L. M. J. & Lüthi, A. Sleep spindles: mechanisms and functions. *Physiol. Rev.* **100**, 805–868 (2020).
6. Brancaccio, A., Tabarelli, D., Bigica, M. & Baldauf, D. Cortical source localization of sleep-stage specific oscillatory activity. *Sci. Rep.* **10**, 6976 (2020).
7. Pauline Dodet, M. H. et al. Sleep disorders in Parkinson's disease, an early and multiple problem. *NPJ Parkinsons Dis.* **10**, 46 (2024).
8. Iber C, A.-I. S., Chesson A., & Quan S. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*, 1st (Westchester, I. L. ed) (American Academy of Sleep Medicine, 2018).
9. Bugalho, P. et al. Polysomnographic predictors of sleep, motor and cognitive dysfunction progression in Parkinson's disease: a longitudinal study. *Sleep. Med.* **77**, 205–208 (2021).
10. Latreille, V. et al. Sleep spindles in Parkinson's disease may predict the development of dementia. *Neurobiol. Aging* **36**, 1083–1090 (2015).
11. Christensen, J. A. et al. Sleep spindle alterations in patients with Parkinson's disease. *Front. Hum. Neurosci.* **9**, 233 (2015).
12. Schreiner, S. J. et al. Reduced regional NREM Sleep slow-wave activity is associated with cognitive impairment in Parkinson disease. *Front. Neurol.* **12**, 618101 (2021).
13. Schreiner, S. J. et al. Slow-wave sleep and motor progression in Parkinson disease. *Ann. Neurol.* **85**, 765–770 (2019).
14. Latreille, V. et al. Electroencephalographic prodromal markers of dementia across conscious states in Parkinson's disease. *Brain* **139**, 1189–1199 (2016).
15. Kalaitzakis, M. E., Gentleman, S. M. & Pearce, R. K. Disturbed sleep in Parkinson's disease: anatomical and pathological correlates. *Neuropathol. Appl. Neurobiol.* **39**, 644–653 (2013).
16. Stephansen, J. B. et al. Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. *Nat. Commun.* **9**, 5229 (2018).
17. Perslev, M. et al. U-Sleep: resilient high-frequency sleep staging. *NPJ Digit Med.* **4**, 72 (2021).
18. Brink-Kjaer, A. et al. Automatic detection of cortical arousals in sleep and their contribution to daytime sleepiness. *Clin. Neurophysiol.* **131**, 1187–1203 (2020).
19. Vallat, R. & Walker, M. P. An open-source, high-performance tool for automated sleep staging. *Elife* **10**, e70092 (2021).
20. Ferri, R. et al. A quantitative statistical analysis of the submentalis muscle EMG amplitude during sleep in normal controls and patients with REM sleep behavior disorder. *J. Sleep. Res.* **17**, 89–100 (2008).
21. Cesari, M. et al. A data-driven system to identify REM sleep behavior disorder and to predict its progression from the prodromal stage in Parkinson's disease. *Sleep. Med.* **77**, 238–248 (2021).
22. Brink-Kjaer, A. et al. End-to-end deep learning of polysomnograms for classification of REM sleep behavior disorder. *Annu Int Conf. IEEE Eng. Med. Biol. Soc.* **2022**, 2941–2944 (2022).

23. Bakker, J. P. et al. Scoring sleep with artificial intelligence enables quantification of sleep stage ambiguity: hypnodensity based on multiple expert scorers and auto-scoring. *Sleep* **46**, zsac154 (2023).
24. Fiorillo, L. et al. U-Sleep's resilience to AASM guidelines. *NPJ Digit Med.* **6**, 33 (2023).
25. Danker-Hopfe, H. et al. Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. *J. Sleep. Res.* **13**, 63–69 (2004).
26. Breen, D. P. et al. Sleep and circadian rhythm regulation in early Parkinson disease. *JAMA Neurol.* **71**, 589–595 (2014).
27. Diederich, N. J., Rufra, O., Pieri, V., Hipp, G. & Vaillant, M. Lack of polysomnographic Non-REM sleep changes in early Parkinson's disease. *Mov. Disord.* **28**, 1443–1446 (2013).
28. Orso, B. et al. Polysomnographic correlates of sleep disturbances in de novo, drug naïve Parkinson's Disease. *Neurol. Sci.* **43**, 2531–2536 (2022).
29. Bolitho, S. J. et al. Disturbances in melatonin secretion and circadian sleep-wake regulation in Parkinson disease. *Sleep. Med.* **15**, 342–347 (2014).
30. Schroeder, L. A. et al. Reduced rapid eye movement density in Parkinson's disease: a polysomnography-based case-control study. *Sleep* **39**, 2133–2139 (2016).
31. Bušková, J. et al. Sleep disturbances in untreated Parkinson's disease. *J. Neurol.* **258**, 2254–2259 (2011).
32. Brunner, H. et al. Microstructure of the non-rapid eye movement sleep electroencephalogram in patients with newly diagnosed Parkinson's disease: effects of dopaminergic treatment. *Mov. Disord.* **17**, 928–933 (2002).
33. Amato, N. et al. Levodopa-induced dyskinesia in Parkinson disease: sleep matters. *Ann. Neurol.* **84**, 905–917 (2018).
34. Margis, R., Schönwald, S. V., Carvalho, D. Z., Gerhardt, G. J. & Rieder, C. R. NREM sleep alpha and sigma activity in Parkinson's disease: evidence for conflicting electrophysiological activity? *Clin. Neurophysiol.* **126**, 951–958 (2015).
35. Christensen, J. A. et al. Decreased sleep spindle density in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. *Clin. Neurophysiol.* **125**, 512–519 (2014).
36. Christensen, J. A. E. et al. Sleep stability and transitions in patients with idiopathic REM sleep behavior disorder and patients with Parkinson's disease. *Clin. Neurophysiol.* **127**, 537–543 (2016).
37. Montini, A. et al. Scoring sleep in neurodegenerative diseases: a pilot study in the synucleinopathies. *Sleep. Med.* **110**, 268–286 (2023).
38. Decat, N. et al. Beyond traditional sleep scoring: massive feature extraction and data-driven clustering of sleep time series. *Sleep. Med.* **98**, 39–52 (2022).
39. Romenets, S. R. et al. Rapid eye movement sleep behavior disorder and subtypes of Parkinson's disease. *Mov. Disord.* **27**, 996–1003 (2012).
40. Rahayel, S. et al. Brain atrophy in Parkinson's disease with polysomnography-confirmed REM sleep behavior disorder. *Sleep* **42**, zsz062 (2019).
41. Rolinski, M. et al. REM sleep behaviour disorder is associated with worse quality of life and other non-motor features in early Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **85**, 560–566 (2014).
42. Postuma, R. B. et al. Rapid eye movement sleep behavior disorder and risk of dementia in Parkinson's disease: a prospective study. *Mov. Disord.* **27**, 720–726 (2012).
43. Liu, Y. et al. Longitudinal changes in parkinson's disease symptoms with and without rapid eye movement sleep behavior disorder: the Oxford Discovery Cohort Study. *Mov. Disord.* **36**, 2821–2832 (2021).
44. Kühn, A. A. et al. Event-related beta desynchronization in human subthalamic nucleus correlates with motor performance. *Brain* **127**, 735–746 (2004).
45. Levy, R. et al. Dependence of subthalamic nucleus oscillations on movement and dopamine in Parkinson's disease. *Brain* **125**, 1196–1209 (2002).
46. Mallet, N. et al. Disrupted dopamine transmission and the emergence of exaggerated beta oscillations in subthalamic nucleus and cerebral cortex. *J. Neurosci.* **28**, 4795–4806 (2008).
47. Brown, P. Bad oscillations in Parkinson's disease. *J. Neural. Transm. Suppl*. **2006**, 27–30 (2006).
48. Neumann, W. J. et al. Long term correlation of subthalamic beta band activity with motor impairment in patients with Parkinson's disease. *Clin. Neurophysiol.* **128**, 2286–2291 (2017).
49. Urrestarazu, E. et al. Beta activity in the subthalamic nucleus during sleep in patients with Parkinson's disease. *Mov. Disord.* **24**, 254–260 (2009).
50. Xie, L. et al. Sleep drives metabolite clearance from the adult brain. *Science* **342**, 373–377 (2013).
51. Nedergaard, M. & Goldman, S. A. Glymphatic failure as a final common pathway to dementia. *Science* **370**, 50–56 (2020).
52. Ju, Y. E., Lucey, B. P. & Holtzman, D. M. Sleep and Alzheimer disease pathology–a bidirectional relationship. *Nat. Rev. Neurol.* **10**, 115–119 (2014).
53. Kang, J. E. et al. Amyloid-beta dynamics are regulated by orexin and the sleep-wake cycle. *Science* **326**, 1005–1007 (2009).
54. Chong, P. L. H., Garic, D., Shen, M. D., Lundgaard, I. & Schwichtenberg, A. J. Sleep, cerebrospinal fluid, and the glymphatic system: a systematic review. *Sleep. Med Rev.* **61**, 101572 (2022).
55. Hablitz, L. M. et al. Increased glymphatic influx is correlated with high EEG delta power and low heart rate in mice under anesthesia. *Sci. Adv.* **5**, eaav5447 (2019).
56. Jiang-Xie, L. F. et al. Neuronal dynamics direct cerebrospinal fluid perfusion and brain clearance. *Nature* **627**, 157–164 (2024).
57. Leng, Y., Musiek, E. S., Hu, K., Cappuccio, F. P. & Yaffe, K. Association between circadian rhythms and neurodegenerative diseases. *Lancet Neurol.* **18**, 307–318 (2019).
58. Sundaram, S. et al. Establishing a framework for neuropathological correlates and glymphatic system functioning in Parkinson's disease. *Neurosci. Biobehav Rev.* **103**, 305–315 (2019).
59. Biondetti, E. et al. Spatiotemporal changes in substantia nigra neuromelanin content in Parkinson's disease. *Brain* **143**, 2757–2770 (2020).
60. Mollenhauer, B. et al. Nonmotor and diagnostic findings in subjects with de novo Parkinson disease of the DeNoPa cohort. *Neurology* **81**, 1226–1234 (2013).
61. Adams, C. et al. Updated Parkinson's disease motor subtypes classification and correlation to cerebrospinal homovanillic acid and 5-hydroxyindoleacetic acid levels. *Clin. Park Relat. Disord.* **8**, 100187 (2023).
62. Stebbins, G. T. et al. How to identify tremor dominant and postural instability/gait difficulty groups with the Movement Disorder Society Unified Parkinson's disease rating scale: comparison with the unified Parkinson's disease rating scale. *Mov. Disord.* **28**, 668–670 (2013).
63. Zampogna, A. et al. Axial impairment and falls in Parkinson's disease: 15 years of subthalamic deep brain stimulation. *NPJ Parkinsons Dis.* **8**, 121 (2022).
64. Brink-Kjær, A. et al. Arousal characteristics in patients with Parkinson's disease and isolated rapid eye movement sleep behavior disorder. *Sleep* **44**, zsab167 (2021).

Acknowledgements

Thank you to the ICEBERG Study contributors: Sara Sambin, Jonas Ihle, Caroline Weill, David Grabli, Florence Cormier-Dequaire, Louise Laure Mariani, Bertrand Degos, Richard Levy, Fanny Pineau, Julie Socha, Eve Benchetrit, Virginie Czernecki, Marie-Alexandrine Glachant, Sophie Rivaud-Pechoux, Elodie Hainque, Samir Bekadar, Alexis Brice, Suzanne Lesage, Fanny Mochel, Farid Ichou, Vincent Perlbarg, Benoit Colsch, Arthur Tenenhaus, Rahul Gaurav, Nadya Pyatigorskaya, Lydia Yahia-Cherif, Romain Valabrègue, Cécile Galléa, Marie-Odile Habert, Dijana Petrovska, Laetitia Jeancolas, Vanessa Brochard, Alizé Chalançon, Carole Dongmo-

Kenfack, Christelle Laganot, Valentine MaheoThis study was supported by grants from Agence Nationale de la Recherche through 'Investissements d'avenir' program (ANR-10-IAIHU-06 and ANR-11-INBS-0006); Fondation EDF; Fondation Planiol; Société Française de Médecine Esthétique and Energipole; Neuratris; Control-PD (Joint Program—Neurodegenerative Disease Research (JPND) COgNiTive propagation in pRodrOmaL Parkinson's disease). The DeNoPa cohort was supported financially by unrestricted grants from Paracelsus-Elena-Klinik, the Michael J. Fox Foundation for Parkinson's Research, Deutsche Parkinson Vereinigung, Parkinson-Fonds Deutschland, and TEVA Pharma. A.B.K. and P.J. were supported by Innovation Foundation Denmark. P.D. was supported by grants from Philippe Foundation and Servier laboratory.

Author contributions

P.D. and A.B.K.: conception, execution, statistical analysis, review, writing on the first draft of the manuscript; E.D., E.M., and P.J.: conception, execution, statistical analysis, review, and critique of the manuscript; I.A., C.T., B.M., F.S.D., E.R., M.V., T.A., S.L., J.C.C., G.M., and S.L.S.: collection, review, and critique of the manuscript. All authors have read and approved the manuscript.

Competing interests

S.L-.S. IA has nothing to report. J-C.C. served in advisory boards for Air Liquide, Biogen, Biophytis, Denali, Ever Pharma, Idorsia, Prevail Therapeutic, Theranexus, and UCB and received grants from Sanofi and The Michael J. Fox Foundation. M.V. received unrestricted grants (Iceberg cohort) from Fondation d'Entreprise EDF, Fondation Saint Michel, the Fondation Thérèse and René Planiol pour l'étude du Cerveau, and unrestricted support for research on Parkinson's disease from Energipole (M. Mallart) and Société Française de Médecine Esthétique (M. Legrand). S.L. received grants from Agence Nationale de la Recherche (ANRMNP 2009, Nucleipark), DHOS-Inserm (2010, Nucleipark), France Parkinson (2008), Ecole Neuroscience de Paris, and "Investissements d'avenir" (grants ANR-10-IAIHU-06 and ANR-11-INBS-0006) during the conduct of the study, and a research grant from Biogen Inc. ER received honorarium for speech from Orkyn, Aguettant, Elivie, and for participating in advisory boards from Merz-Pharma, Ipsen, Elivie, Teva, and BIAL. He received research support from Merz-Pharma, Orkyn, Aguettant, Elivie, Ipsen, Everpharma, Enjoysharing, Fondation Desmarest, AMADYS, ADCY5.org, Fonds de dotation Patrick Brou de Laurière, Agence Nationale de la Recherche, Dystonia Medical Research Foundation.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41531-025-01105-w.

Correspondence and requests for materials should be addressed to Pauline Dodet or Andreas Brink-Kjaer.