Supplementary Text

Statistical analysis

Analyses from baseline

For our main analyses, we used Cox proportional hazard models to determine the association of sleep quality, measured with the global PSQI score, and sleep duration at baseline with incident parkinsonism and Parkinson's disease, using follow-up time as timescale. We repeated the main analysis after categorizing sleep quality and duration (sleep duration according to international recommendations for elderly individuals (Hirshkowitz *et al.*, 2015)).

When investigating the assumption of proportional hazards of the Cox model through visually examining and statistically testing the scaled *Schoenfeld* residuals, we found slight ($0.01<p<0.05$) violations of proportionality for both sleep determinants in the main analyses on Parkinson's disease. This indicates that the hazard ratio, which provides an average of the relative risk over the included follow-up time, is a poor representation of the changes occurring over time within the study timeframe (Hernan, 2010). To examine these changes over time, we repeated the main analyses for both outcomes after restricting follow-up to shorter study duration by censoring participants at 2, 4, 6, 8, and 10 years after baseline, using Firth's penalized Cox regression analysis to account for low cumulative incidences of outcomes after short follow-up (Firth, 1993; Heinze and Dunkler, 2008). Such an approach shows how the choice of follow-up time from baseline affects the hazard ratio (Hernan, 2010). Of note, proportionality was not violated for analyses of 6 years after baseline or shorter (for the association of sleep duration and Parkinson's disease), or at 2 years after baseline (for the association of sleep quality and Parkinson's disease). To further examine hazard ratio changes over time, we used a stratified Cox model by stratifying by follow-up time intervals of 0-2, 2-4, 4-6, 6-8, 8-10, and 10-13 years. Hazard ratios were obtained by

modeling the interaction of the determinant with a term of categorized follow-up time, and combining the coefficients of point and interval estimation for the determinant and that stratum.

Next, to investigate if any associations found for global PSQI score were driven by specific components, we also investigated the relation between PSQI component scores (quality, latency, efficiency, disturbances, medication and daytime dysfunction) and incident parkinsonism or Parkinson's disease, in overall and shorter follow-up durations similarly as described above. As sleep duration was already investigated separately, we did not additionally investigate the duration component of the PSQI (which categorizes reported sleep duration (Buysse *et al.*, 1989)). Also, we performed the main analyses in persons without any comorbid clinically relevant depressive symptoms (CES-D score ≥16) and without any anxiety disorders. Furthermore, we studied possible effect modification in the main analyses by median age, sex, and presence versus absence of any of four parkinsonian signs (scoring details published previously (Darweesh *et al.*, 2017)), by performing stratified analyses and formally testing for multiplicative interaction. Proportionality was not violated in the tests of multiplicative interaction.

Change analyses between the baseline and follow-up visit

We also examined how changes in sleep quality and duration between the baseline and follow-up visit were related to subsequent risk of parkinsonism and Parkinson's disease. Follow-up time was calculated from the follow-up visit, and analyses were additionally adjusted for the time interval between the baseline and follow-up visit. A change in sleep quality was modeled by subtracting the baseline global PQSI score from the score at the follow-up visit, so that positive values indicated worsening of sleep quality over time. Change in sleep duration was modeled as shorter sleep duration, by subtracting self-reported sleep

duration at the follow-up visit from that at baseline. We repeated analyses after i) additionally adjusting for averaged global PSQI score, or sleep duration, over baseline and follow-up visits to examine if effects were dependent on absolute levels (i.e. if decreases in e.g. sleep duration from 9 to 7 hours would be different from decreases from 7 to 5 hours); ii) adjusting for depressive symptoms at baseline to see if sleep changes were driven by depression. As we also observed non-proportionality of hazard ratios ($0.01<p<0.05$) in the analyses of changes in sleep quality and duration between the baseline and follow-up visit on the risk of parkinsonism, we also obtained period-specific hazard ratios for these relations.

To obtain normally distributed values and minimize the effect of outliers, we log-transformed (ln(variable + 1)) right-skewed variables (global PSQI score) and subsequently winsorized (i.e. transformed towards the mean) outliers to three standard deviations from the mean (1.3% of observations for global PSQI score, 0.8% for sleep duration). Both variables were then standardized (subtracting the mean and dividing by the standard deviation) to facilitate comparison of effect sizes.

Missing data on covariates (missing values in covariates at baseline: median=1.6%, maximum=29.7% (smoking status)) were imputed using five multiple imputation based on all variables used in our analyses. Statistical testing was performed two-sided at $p<0.05$. Data were analysed using *SPSS Statistics*, version 21 (IBM Corp., Armonk, NY), and with the open *R* software (packages: 'survival', 'coxphf').

Supplementary references

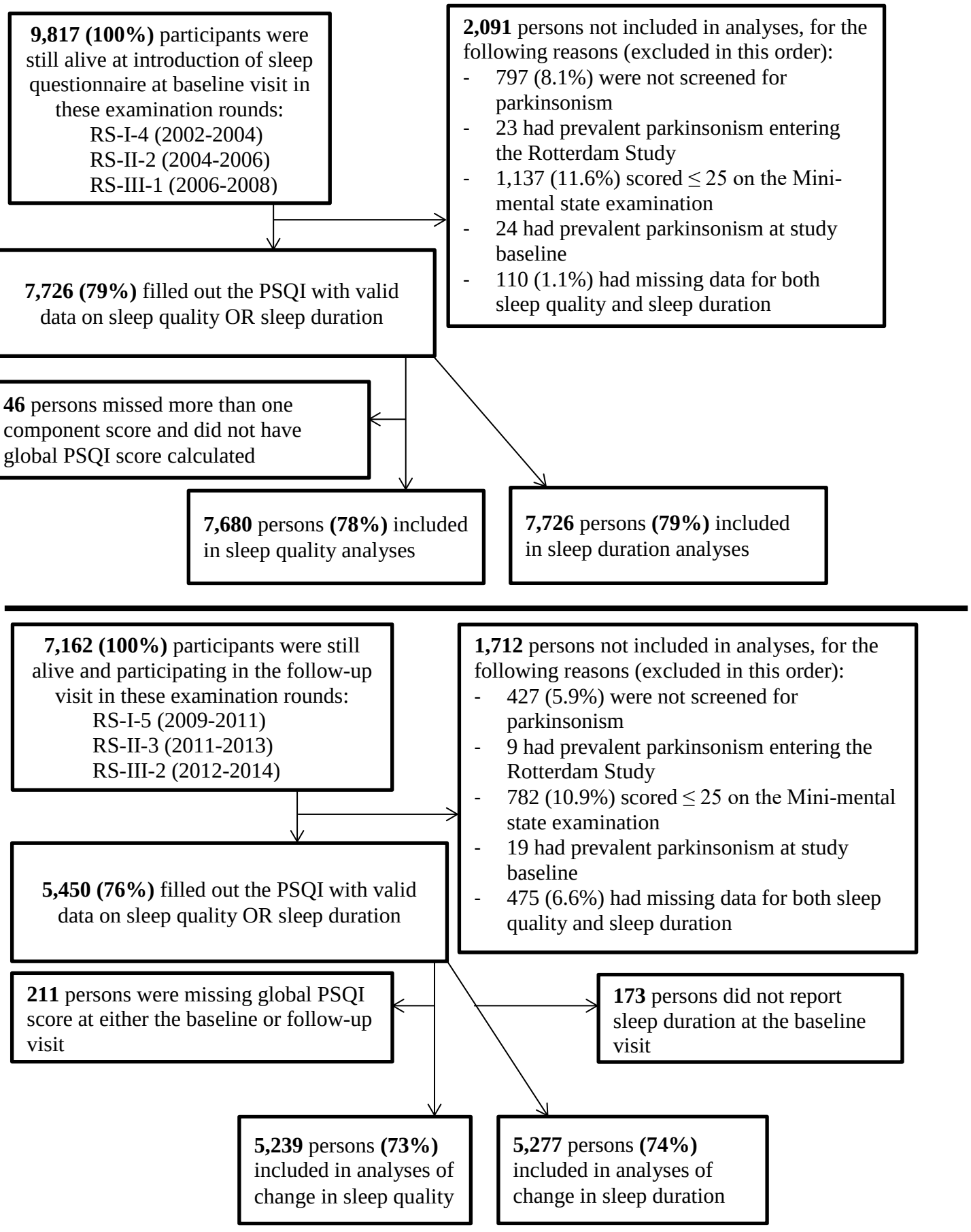
Buysse DJ, Reynolds CF, 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989; 28(2): 193-213.

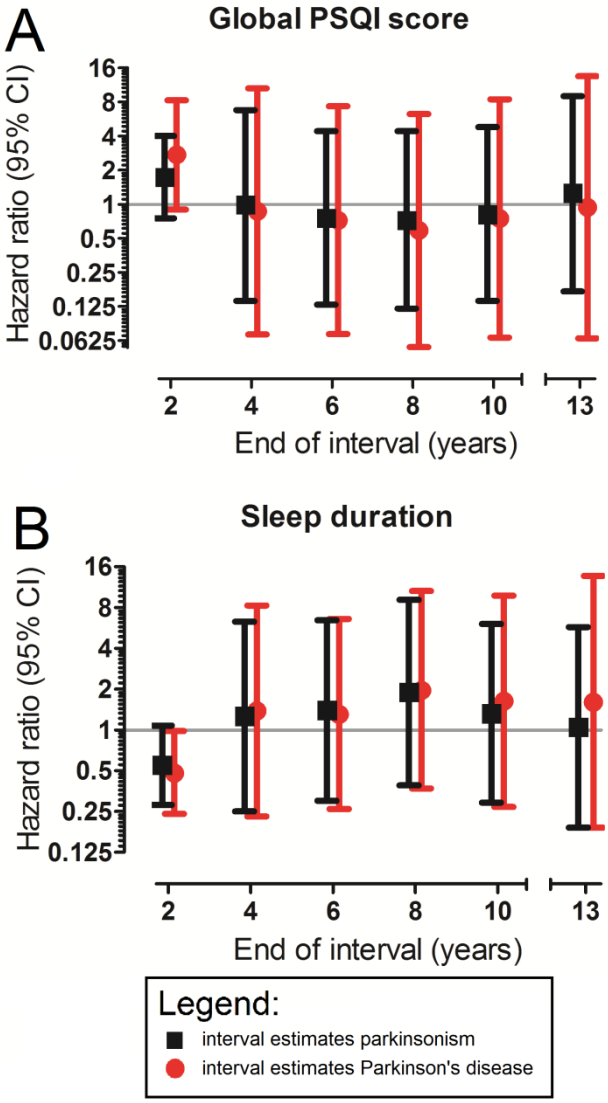
Darweesh SK, Verlinden VJ, Stricker BH, Hofman A, Koudstaal PJ, Ikram MA. Trajectories of prediagnostic functioning in Parkinson's disease. Brain 2017; 140(2): 429-41.

Firth D. Bias reduction of maximum likelihood estimates. Biometrika 1993; 80: 27-83.

Heinze G, Dunkler D. Avoiding infinite estimates of time-dependent effects in small-sample survival studies. Stat Med 2008; 27(30): 6455-69.
Hernan MA. The hazards of hazard ratios. Epidemiology 2010; 21(1): 13-5.
Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, *et al.* National Sleep Foundation's updated sleep duration recommendations: final report. Sleep Health 2015; 1(4): 233-43.

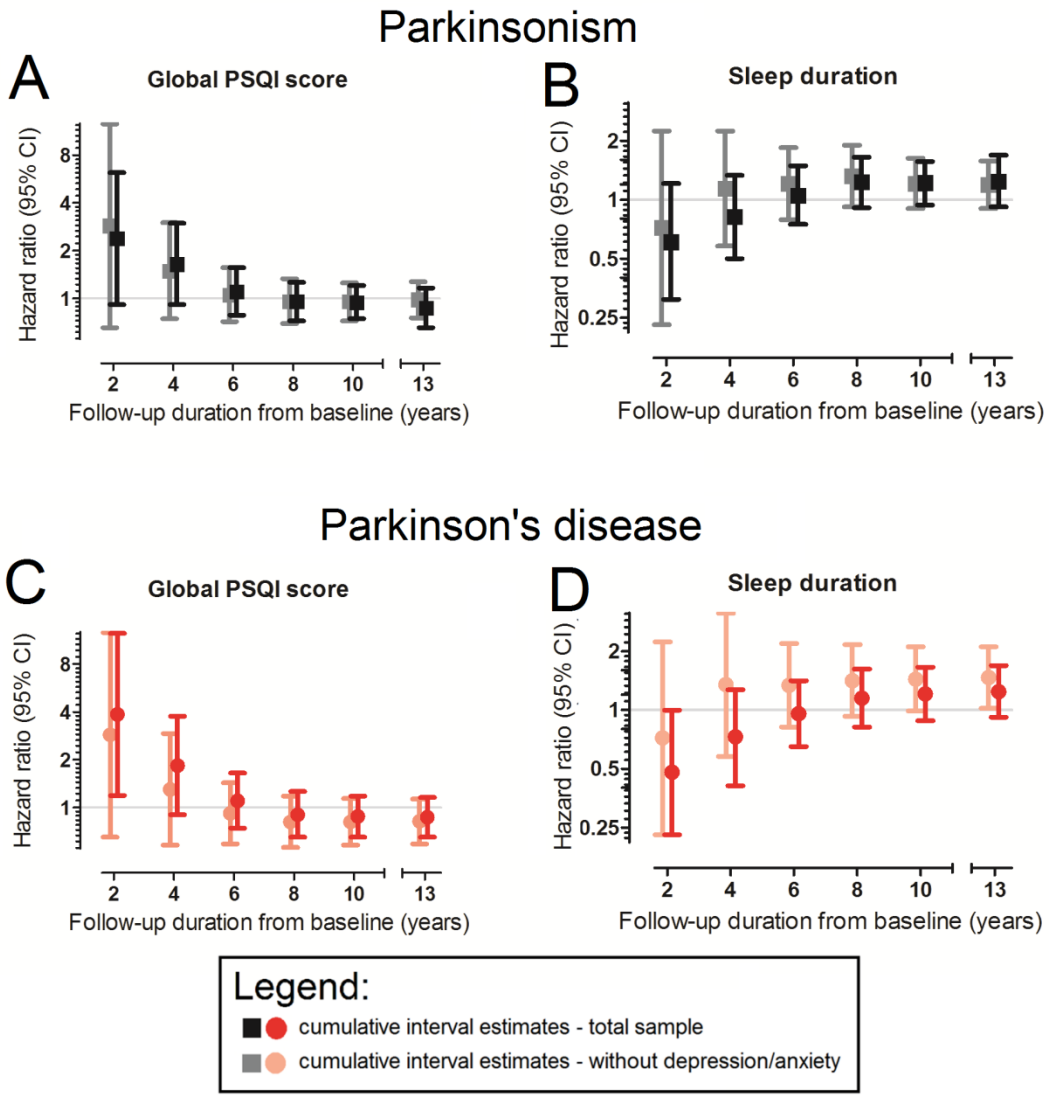
Supplementary Figure 1: Flowchart inclusion of study participants at the baseline and follow-up visits





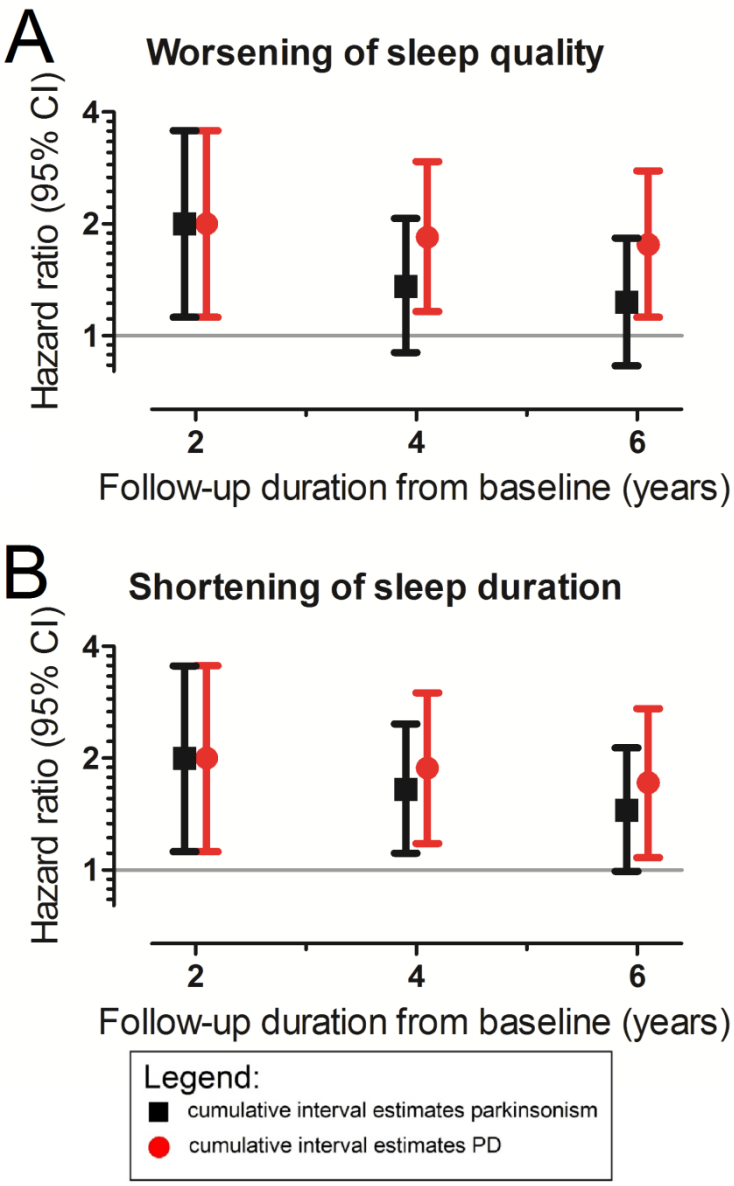
Supplementary Figure 2: Associations of sleep quality and duration with risk of parkinsonism and Parkinson's disease, over separate intervals of follow-up time

The associations of (A) sleep quality and (B) sleep duration with incident parkinsonism and Parkinson's disease are shown for separate intervals of follow-up duration within the study timeframe, using a stratified Cox model. Hazard ratio were estimated for the intervals 0-2 years, 2-4 years, 4-6 years, 6-8 years, 8-10 years, and 10-13 years (end of follow-up) and obtained from modeling the interaction of the determinants with follow-up time strata. Hazard ratio estimates were adjusted for age at baseline, sex, educational level and smoking status, are expressed per standard deviation increase of (A) worse sleep quality, or (B) longer sleep duration, and are plotted at a logarithmic (base 2) scale. Abbreviations: CI=Confidence Interval; PD=Parkinson's disease; PSQI=Pittsburgh Sleep Quality Index



Supplementary Figure 3: Associations of sleep quality and duration with parkinsonism and Parkinson's disease in persons without comorbid depression and anxiety, per cumulatively increasing duration of follow-up

Associations of (A) sleep quality and (B) sleep duration with incident parkinsonism, and (C) sleep quality and (D) sleep duration with incident Parkinson's disease, analyzed both in the total sample and in persons without clinically relevant depressive symptoms and anxiety disorders. Associations are depicted for cumulatively increasing follow-up duration within the study timeframe. Hazard ratio estimates were obtained from multivariate Firth's penalized Cox regression models by censoring all participants still at risk at year 2, 4, 6, 8 and 10 after baseline, and after the total follow-up of 13 years. Estimates were adjusted for age at baseline, sex, educational level and smoking status, are expressed per standard deviation increase of PSQI score or sleep duration, and are plotted at a logarithmic (base 2) scale. Abbreviations: CI=Confidence Interval; PSQI=Pittsburgh Sleep Quality Index



Supplementary Figure 4: Associations of changes in sleep quality and duration between the baseline and follow-up visit with risk of parkinsonism and Parkinson's disease, per cumulatively increasing duration of follow-up

The associations of changes in (A) sleep quality ('worsening') and (B) sleep duration ('shortening') between the baseline and follow-up visit with incident parkinsonism and Parkinson's disease are shown for cumulatively increasing follow-up duration within the six-year study timeframe. Hazard ratio estimates were obtained from multivariate Firth's penalized Cox regression models by censoring all participants still at risk at year 2, 4, and after the total follow-up of 6 years. Hazard ratio estimates were adjusted for age at baseline, sex, educational level, smoking status, and time interval between measurements, are expressed per standard deviation increase of (A) worsening sleep quality, or (B) shorter sleep duration, and are plotted at a logarithmic (base 2) scale.

Abbreviations: CI=Confidence Interval; PD=Parkinson's disease; PSQI=Pittsburgh Sleep Quality Index

Supplementary Table 1: Overview of incident parkinsonism diagnoses

Clinical diagnosis	After the baseline visit (N=7,726)	After the follow-up visit (N=5,450)
Probable Parkinson's disease	47 (63%)	17 (68%)
Vascular parkinsonism	3 (4%)	1 (4%)
Medication-induced parkinsonism	5 (7%)	1 (4%)
Progressive supra-nuclear palsy	1 (1%)	0
Multiple system atrophy	0	0
Corticobasal degeneration	1 (1%)	1 (4%)
Lewy body dementia	2 (3%)	0
Parkinsonism with dementia – not Lewy body type	2 (3%)	0
Unspecified parkinsonism*	14 (19%)	5 (20%)
All parkinsonism diagnoses	**75 (100%)**	**25 (100%)**

Number of diagnoses expressed as frequency (%), for the samples used in analyses at the baseline and follow-up visits.

*Denotes parkinsonism patients that did not have any of the above clinical diagnoses

Supplementary Table 2: Association between categories of sleep quality or duration and risk of incident parkinsonism or Parkinson's disease

Determinant	Categories	Parkinsonism		Parkinson's disease	
		Cases/N	HR (95% CI)	Cases/N	HR (95% CI)
Global PSQI score	≤5 ('good' quality)	55/5,565	1.00 (reference)	36/5,562	1.00 (reference)
	>5 ('poor' quality)	20/2,115	0.97 (0.57-1.66)	11/2,112	0.79 (0.39-1.59)
Sleep duration	<7 hours	21/3,155	1.00 (reference)	13/3,150	1.00 (reference)
	≥7 - ≤8 hours	45/4,033	1.61 (0.95-2.71)	28/4,031	1.65 (0.86-3.21)
	>8 hours	9/538	2.19 (1.00-4.81)	6/537	2.54 (0.96-6.72)

Hazard ratios were obtained from Cox regression models, adjusted for age at baseline, sex, educational level, and smoking status, expressed in reference to the lowest global PSQI score, or sleep duration, category. Categorization of sleep duration in three categories is based on the US National Sleep Foundation recommended sleep duration for elderly persons (Hirshkowitz *et al.*, 2015).

Abbreviations: HR=Hazard ratio; N=sample size; PSQI=Pittsburgh Sleep Quality Index.

Supplementary Table 3: Associations of Pittsburgh Sleep Quality Index component scores with risk of parkinsonism and Parkinson's disease, per cumulatively increasing duration of follow-up

PSQI component	Outcome	N	Duration of follow-up time in years					
			≤2	≤4	≤6	≤8	≤10	Overall
			Hazard ratio per point increase (worse sleep) on the component score (95% CI)					
Quality	PS	7716	3.16 (0.95; 10.46)	**2.33 (1.05; 5.17)**	1.16 (0.68; 1.98)	1.08 (0.69; 1.69)	1.07 (0.73; 1.56)	1.09 (0.76; 1.57)
	PD	7709	**3.90 (1.04; 14.58)**	**2.57 (1.01; 6.57)**	1.17 (0.63; 2.18)	1.01 (0.60; 1.70)	0.93 (0.57; 1.50)	0.95 (0.60; 1.50)
Latency	PS	7718	2.48 (0.90; 6.83)	**2.27 (1.14; 4.51)**	1.54 (0.99; 2.40)	1.27 (0.87; 1.86)	1.11 (0.80; 1.55)	1.15 (0.84; 1.57)
	PD	7712	2.94 (0.97; 8.87)	2.16 (0.97; 4.79)	1.53 (0.91; 2.56)	1.19 (0.76; 1.85)	1.05 (0.70; 1.59)	1.06 (0.71; 1.57)
Efficiency	PS	7473	2.78 (0.97; 7.98)	**1.98 (1.04; 3.77)**	1.42 (0.90; 2.24)	1.05 (0.68; 1.60)	1.05 (0.73; 1.51)	0.99 (0.70; 1.41)
	PD	7466	**4.54 (1.27; 16.18)**	**2.35 (1.09; 5.07)**	1.65 (1.00; 2.73)	1.17 (0.73; 1.88)	1.09 (0.69; 1.71)	1.00 (0.64; 1.56)
Disturbances	PS	6840	0.70 (0.15; 3.25)	0.85 (0.29; 2.44)	1.08 (0.53; 2.21)	0.94 (0.52; 1.69)	0.75 (0.45; 1.23)	0.78 (0.48; 1.26)
	PD	6835	1.08 (0.20; 6.01)	0.96 (0.27; 3.34)	0.96 (0.42; 2.21)	0.76 (0.38; 1.52)	0.64 (0.34; 1.19)	0.64 (0.35; 1.18)
Sleep medication	PS	7725	1.08 (0.14; 8.32)	0.54 (0.09; 3.23)	1.28 (0.49; 3.33)	1.00 (0.42; 2.37)	1.29 (0.65; 2.56)	1.08 (0.55; 2.14)
	PD	7718	1.54 (0.19; 12.76)	0.73 (0.11; 4.68)	1.04 (0.32; 3.41)	0.95 (0.34; 2.66)	1.19 (0.50; 2.85)	1.06 (0.45; 2.52)
Daytime dysfunction	PS	7689	2.49 (0.48; 12.81)	**3.00 (1.01; 8.89)**	1.64 (0.72; 3.74)	1.40 (0.68; 2.87)	1.10 (0.56; 2.14)	0.94 (0.48; 1.82)
	PD	7684	3.34 (0.60; 18.55)	2.31 (0.62; 8.71)	1.24 (0.44; 3.51)	1.04 (0.41; 2.63)	0.85 (0.34; 2.10)	0.76 (0.31; 1.89)

The associations of the PSQI components with incident parkinsonism and Parkinson's disease are provided for cumulatively increasing follow-up duration within the study timeframe. Hazard ratio estimates were obtained from multivariate Firth's penalized Cox regression models by censoring all participants still at risk at year 2, 4, 6, 8 and 10 after baseline, and after the total follow-up of 13 years. Estimates are adjusted for age at baseline, sex, educational level, and smoking status, and are expressed per category increase in component score. To ensure sufficient (>10%) observations in each category, we combined scores 2 and 3 for components quality, latency and efficiency, and scores 1, 2 and 3 for components disturbances, medication and daytime dysfunction.

Abbreviations: CI=Confidence interval; N=Sample size; PD=Parkinson's disease; PS=parkinsonism; PSQI=Pittsburgh Sleep Quality Index.

Supplementary Table 4: Association of sleep quality or duration and risk of incident Parkinson's disease or parkinsonism, stratified by potential effect-modifiers

Effect-modifier	Strata	Cases/N	Parkinsonism HR per SD increase (95% CI)	P_{int}	Cases/N	Parkinson's dis. HR per SD increase (95% CI)	P_{int}
Sleep quality							
Age[a]	≤ 67.5 years	12/384	0.56 (0.31 – 1.00)	0.816	10/3837	0.62 (0.32 - 1.19)	0.683
	> 67.5 years	63/3840	1.07 (0.82 - 1.38)		37/3837	0.95 (0.69 - 1.32)	
Sex	Male	42/3305	1.01 (0.74 - 1.38)	0.289	24/3300	0.81 (0.53 - 1.22)	0.479
	Female	33/4375	0.89 (0.63 - 1.25)		23/4374	0.93 (0.62 - 1.41)	
Parkinsonian	Present	16/804	**1.96 (1.07 - 3.59)**	**0.004**	12/802	1.53 (0.80 - 2.91)	**0.048**
signs	Absent	59/6876	0.80 (0.62 - 1.03)		35/6872	**0.72 (0.52 – 1.00)**	
Sleep duration							
Age[a]	≤ 67.5 years	12/3863	1.84 (0.97 - 3.50)	0.778	10/3859	1.68 (0.82 - 3.47)	0.406
	> 67.5 years	63/3863	1.11 (0.86 - 1.44)		37/3859	1.14 (0.82 - 1.60)	
Sex	Male	42/3330	1.02 (0.73 - 1.43)	0.218	24/3323	1.26 (0.79 - 2.01)	0.870
	Female	33/4396	1.39 (0.98 - 1.97)		23/4395	1.21 (0.80 - 1.82)	
Parkinsonian	Present	16/807	1.00 (0.62 - 1.60)	0.270	12/805	1.09 (0.64 - 1.86)	0.460
signs	Absent	59/6919	1.29 (0.98 - 1.71)		35/6913	1.31 (0.91 - 1.89)	

The associations of sleep quality and sleep duration with incident parkinsonism and Parkinson's disease are shown stratified for several potential effect-modifiers. Hazard ratios were obtained from Cox regression models, adjusted for (if applicable) age at baseline, sex, educational level, and smoking status, and are expressed per standard deviation increase of global Pittsburgh Sleep Quality Index score or sleep duration. Multiplicative interaction was tested in a model including the main effects of the stratified variable, and a untransformed and non-standardized variable of sleep quality or sleep duration.

Bold indicates statistical significance at $P<0.05$.

[a]Split at median age in sample

Abbreviations: CI=Confidence interval; dis=disease; HR=Hazard ratio; N=sample size; P_{int}=P-value interaction term; SD=standard deviation.

Supplementary Table 5: Characteristics of study population at follow-up visit

Characteristic (unit)	N = 5,450
Age at baseline (years)	68.4 ± 8.9
Female	3,127 (57%)
Educational level	
Primary education	398 (7%)
Lower/intermediate or lower vocational	2,161 (39%)
Higher or intermediate vocational	1,636 (30%)
Higher vocational or university	1,259 (23%)
Smoking status	
Never smoker	1,748 (32%)
Former smoker	3,068 (56%)
Current smoker	637 (12%)
Cognitive functioning (MMSE score)	29 (27-29)
Presence of any parkinsonian signs	806 (15%)
Time interval between baseline and follow-up visits (years)	6.0 ± 0.6
Missing	49 (1%)
Sleep quality (global PSQI score)	3 (1-6)
Missing	9 (0%)
Change in sleep quality compared to baseline (global PSQI score increase)	0.0 ± 3.1
Missing	211 (4%)
Sleep duration (hours)	6.9 ± 1.3
Change in sleep duration compared to baseline (hours decrease)	-0.1 ± 1.16
Missing	173 (3%)

Characteristics for eligible study population for analyses of sleep change at the follow-up visit. Values are expressed as frequency (%) for categorical variables and mean ± standard deviation or median (interquartile range) for continuous variables, unless specified otherwise. Includes imputed values for covariates.

Abbreviations: MMSE=Mini-mental state examination; PSQI=Pittsburgh Sleep Quality Index.