npj | dementia

Article

https://doi.org/10.1038/s44400-025-00008-0

# Sleep disturbances as risk factors for neurodegeneration later in life

Check for updates

Emily Simmonds[1,10], Kristin S. Levine[2,3,10], Jun Han[4], Hirotaka Iwaki[2,3], Mathew J. Koretsky[2], Nicole Kuznetsov[2], Faraz Faghri[2,3], Caroline Warly Solsberg[2,3,5,6], Artur Schuh[7,8], Lietsel Jones[2,3], Sara Bandres-Ciga[2], Cornelis Blauwendraat[2], Andrew Singleton[2], Valentina Escott-Price[1,4] ✉, Hampton L. Leonard[2,3,9] ✉ & Mike A. Nalls[2,3]

The relationship between sleep disorders and neurodegeneration is complex. Using >1 million electronic health records from Wales, UK, and Finland, we mined biobank data to identify relationships between sleep disorders and neurodegenerative diseases (NDDs). Additionally, we investigated how sleep-attributed risk may compensate for lack of NDD genetic risk. We found that sleep disorders were associated with risk of Alzheimer's disease (AD), amyotrophic lateral sclerosis, dementia, Parkinson's disease (PD), and vascular dementia in three national biobanks (hazard ratios from 1.3 (PD) to 5.11 (dementia)). Sleep disorders imparted risk up to 15 years before NDD onset. Sleep factors were independent of AD and PD genetic risk, potentially compensating for low genetic risk in disease etiology. Poor sleep hygiene and sleep apnea have several available treatments that could potentially reduce the risk of neurodegeneration. Sleep-related risk factors are significantly and independently enriched in NDD patients with low genetic risk.

The World Health Organization has recognized sleep as a critical health state and health-related behavior[1] [Accessed 30 Aug 2023]. One-fourth of Europeans have insomnia[2]. People with sleep disturbances that prevent them from getting adequate rest commonly have short-term daytime cognitive impairment and individuals who experience sleep problems have been shown to have a greater risk of developing dementia[3,4].

Although an association between sleep and neurological disorders is widely acknowledged, it is largely unclear how dysfunctions in sleep and circadian rhythms contribute to the etiology of neurodegeneration or whether they are a contributing cause or consequence of these conditions. Many individuals who develop dementia experience sleep problems following dementia onset, and there is evidence that these conditions are involved in a complex self-reinforcing bidirectional relationship (e.g., relationship between amyloid-β plaque accumulation and poor sleep)[5]. As dysregulation of the circadian clock already occurs during the asymptomatic stage of the disease and could promote neurodegeneration, restoration of sleep and circadian rhythms in preclinical neurodegenerative disorders may represent an opportunity for early intervention to slow the disease course.

Understanding the genetic interaction between sleep disorders and neurodegeneration is crucial to understanding if sleep disorders are merely an early sign of an NDD or if they are risk factors that affect the course of disease. If sleep disorders are risk factors, then appropriately treating them could potentially lower the risk of developing an NDD and/or delay age of onset.

Sleep disturbances have been associated with future risk of both cognitive impairment and AD pathology, and can be an early indication of presymptomatic stages of AD[6,7]. Various types of dementia are associated with different types of sleep and circadian disturbances[6–8]. For example, obstructive sleep apnea, a primary sleep disorder marked by nighttime airway collapse leading to brief lapses in breathing, is associated with dementia in general and is particularly common with Alzheimer's disease, occurring in about 50% of patients[9]. Patients with insomnia also have a higher risk of dementia[10] and the prevalence and severity of sleep disorders increase with dementia severity[11]. Sleep disturbance can occur very early in AD or even precede the development of cognitive problems[12]; the preclinical stage of AD has been associated with worse sleep quality and shorter sleep duration[13].

[1]UK Dementia Research Institute (UKDRI) at Cardiff University, Cardiff, UK. [2]Center for Alzheimer's and Related Dementias, National Institutes of Health, Bethesda, MD, USA. [3]DataTecnica LLC, Washington, DC, USA. [4]Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK. [5]Pharmaceutical Sciences and Pharmacogenomics, UCSF, San Francisco, CA, USA. [6]Department of Neurology and Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, USA. [7]Departamento de Farmacologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil. [8]Serviço de Neurologia, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil. [9]German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany. [10]These authors contributed equally: Emily Simmonds, Kristin S. Levine. ✉e-mail: EscottPriceV@cardiff.ac.uk; hampton@datatecnica.com

PD and sleep are connected in complex ways that are not completely understood. Sleep-related symptoms may be one of the earliest signs of Parkinson's disease[14]. The majority of patients with rapid-eye movement sleep behavior disorder (RBD) eventually develop PD or another neurological condition[15]. A recent study found risk loci for RBD near known PD genes, such as SNCA and GBA[16]. Other sleep disorders noted in individuals with PD include insomnia, excessive daytime sleepiness, restless leg syndrome and sleep disordered breathing with insomnia being the most common[17].

Sleep disorders have also been reported with other neurodegenerative disorders such as vascular dementia, where sleep disturbances and daytime sleepiness have been shown to predict vascular dementia[18], multiple sclerosis, where sleep disorders, fatigue and daytime sleepiness are commonly reported symptoms[19], and amyotrophic lateral sclerosis (ALS) which is linked to sleep disorders often related to physical muscular symptoms[20].

To unravel the causal structure that relates to NDDs and sleep disorders, it is helpful to examine individuals over as long a time period as possible, as well as differentiate between individuals who were diagnosed with sleep disorders pre- and post-NDD diagnosis. For this, medical health records are an invaluable tool, as they provide time-stamped information on all medical events, including symptoms, diagnoses, medication usage, and clinical interventions across an individual's full-time while enrolled in that system. Here, we use three national scale biobanks with massive sample sizes to: look for associations between ICD10 sleep disorder codes and NDD diagnosis; conduct a lag analysis to look at first exposure at different points in time; adjust for the number of sleep disorder diagnoses; and explore the interaction between sleep disorders and genetic risk of NDD (polygenic risk scores) in order to disentangle the complexities of sleep and neurodegeneration.

## Results

### Prior sleep disorders are associated with late-life neurodegeneration

The pairs of sleep disorders and NDDs shown to have significant associations in the meta analysis are summarized in Table 1. The associations in the individual cohorts are available in Supplementary Table 2. N_pairs indicates the number of individuals with both an NDD diagnosis and a sleep disorder. AD, dementia, vascular dementia, and PD were all significantly associated with the code G47 which encompasses sleep disorders associated with circadian rhythm, such as narcolepsy, apnea, hyper and parasomnia, as well as cataplexy and movement-related sleep issues. G47 coding associated with these diseases exhibited hazard ratios (HRs) ranging from 1.15 (95% CI 1.09–1.22 for AD) to 1.41 (95% CI 1.29–1.54 for vascular dementia). For a graphical summary of these results, see Fig. 1.

Non-organic sleep disorders under ICD10 code F51 include mental, behavioral, and neurodevelopmental conditions such as nightmares and generalized insomnia (without substance abuse). They are significantly associated with increased risk of dementia, PD, and vascular dementia. These associations ranged in HRs from 1.67 (95% CI 1.42–1.97 in dementia) to 2.05 (95% CI 1.60–2.62 in vascular dementia). In general, these are the largest HRs observed in this report.

We also investigated whether the granular subcategories were associated with NDDs. Sleep apnea (ICD10 code G47.3) specifically was associated with dementia (HR 1.34, 95% CI 1.24–1.45) and vascular dementia (HR 1.44, 95% CI 1.35–1.54), suggesting that there could be cardiometabolic connections to risk.

Eight pairs of sleep disorders and NDDs were shown to have significant associations in two datasets (see Table 1). Sleep apnea was significant for ALS (HR 1.47, 95% CI 1.31–1.65) and PD (HR 1.17, 95% CI 1.01–1.36) in the meta-analysis. These associations for ALS should be taken with caution as ALS is relatively rare across biobanks; no ALS data was available in SAIL. For PD, G47.8 (HR 2.43, 95% CI 1.99–2.98) and G47.9 (HR 1.47, 95% CI 1.27–1.71) were also significantly associated with the meta-analysis. G47.8 and G47.9 are both "other" sleep disorders codes. Since REM sleep disorder was not given an ICD10 code until 2016, it is possible that some of these codings may reflect that disorder.

We tested the proportional hazards assumption for all models, and no violation was detected. We also used models where sleep disorders were included as a time-dependent covariate; the results did not change substantially so we chose the simpler model for this paper.

Since sleep apnea was the sleep code most associated with NDDs, we also investigated the impact of individuals using a continuous positive airway pressure (CPAP) machine. Surprisingly, in SAIL, a low percentage of individuals using a CPAP machine are also recorded as having sleep apnea (~5–15%). It may be that recording the use of a CPAP machine in a clinic is used as a proxy to sleep apnea diagnosis. Cox models in the SAIL data using whether or not an individual used a CPAP machine as the independent variable, show almost identical hazard ratios to those using sleep apnea (G47.3), and when both variables are included into the model, the effect of CPAP is removed, due to the collinearity between sleep apnea and CPAP. This is consistent across all NDDs studied here. Due to the similar effect sizes for sleep apnea and CPAP, it is likely that the results will be consistent if

**Table 1 | Sleep/NDD meta-analysis results**

| ICD10 | PRIOR | OUTCOME | metaHR | ci_min | ci_max | P_VAL | N_SIG_COHORTS* |
|---|---|---|---|---|---|---|---|
| G47 | Other sleep disorders | AD | 1.15 | 1.09 | 1.22 | <1.0E-04 | 3 |
| F51 | Non-organic sleep disorders | DEMENTIA | 1.67 | 1.42 | 1.97 | <1.0E-04 | 3 |
| G47 | Other sleep disorders | DEMENTIA | 1.35 | 1.22 | 1.49 | <1.0E-04 | 3 |
| G47.3 | Sleep apnea | DEMENTIA | 1.34 | 1.24 | 1.45 | <1.0E-04 | 3 |
| G47 | Other sleep disorders | PD | 1.24 | 1.04 | 1.49 | 1.8E-02 | 3 |
| G47.3 | Sleep apnea | VASCULAR | 1.44 | 1.35 | 1.54 | <1.0E-04 | 3 |
| F51 | Nonorganic sleep disorders | PD | 1.68 | 1.33 | 2.13 | <1.0E-04 | 2 |
| F51 | Nonorganic sleep disorders | VASCULAR | 2.05 | 1.6 | 2.62 | <1.0E-04 | 2 |
| G47 | Other sleep disorders | VASCULAR | 1.41 | 1.29 | 1.54 | <1.0E-04 | 2 |
| G47.3 | Sleep apnea | ALS | 1.47 | 1.31 | 1.65 | <1.0E-04 | 2 |
| G47.3 | Sleep apnea | PD | 1.17 | 1.01 | 1.36 | 3.5E-02 | 2 |
| G47.8 | Other sleep disorders | PD | 2.43 | 1.99 | 2.98 | <1.0E-04 | 2 |
| G47.9 | sleep disorder, unspecified | DEMENTIA | 1.44 | 1.32 | 1.57 | <1.0E-04 | 2 |
| G47.9 | sleep disorder, unspecified | PD | 1.47 | 1.27 | 1.71 | <1.0E-04 | 2 |

*N_SIG_COHORTS is the number of significant associations in the three independent cohorts.

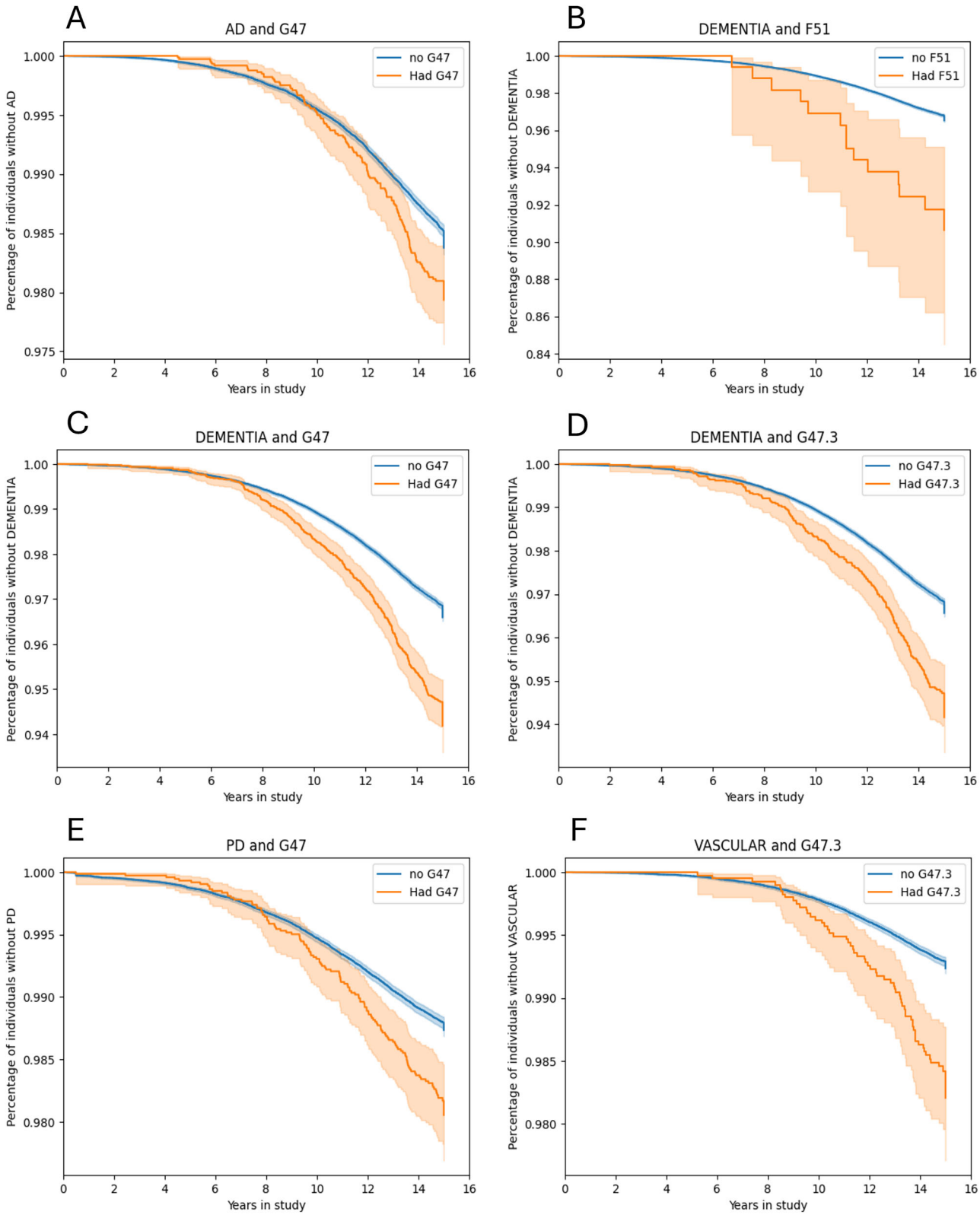


**Fig. 1 | Kaplan-Meier Plots for Replicated NDD/Sleep code pairs showing longitudinal differences in disease risk between individuals with sleep disturbances and those without.** Only replicated pairings are plotted. **A** AD and G47; **B** Dementia and F51; **C** Dementia and G47; **D** Dementia and G47.3; **E** PD and G47; **F** Vascular Dementia and G47.3. G47: Sleep disorders associated with circadian rhythm, such as narcolepsy, apnea, hyper and parasomnia, as well as cataplexy and movement-related sleep issues. **F51**: Non-organic sleep disorders, such as nightmares and generalized insomnia (without substance abuse). **G47.3**: Sleep apnea.

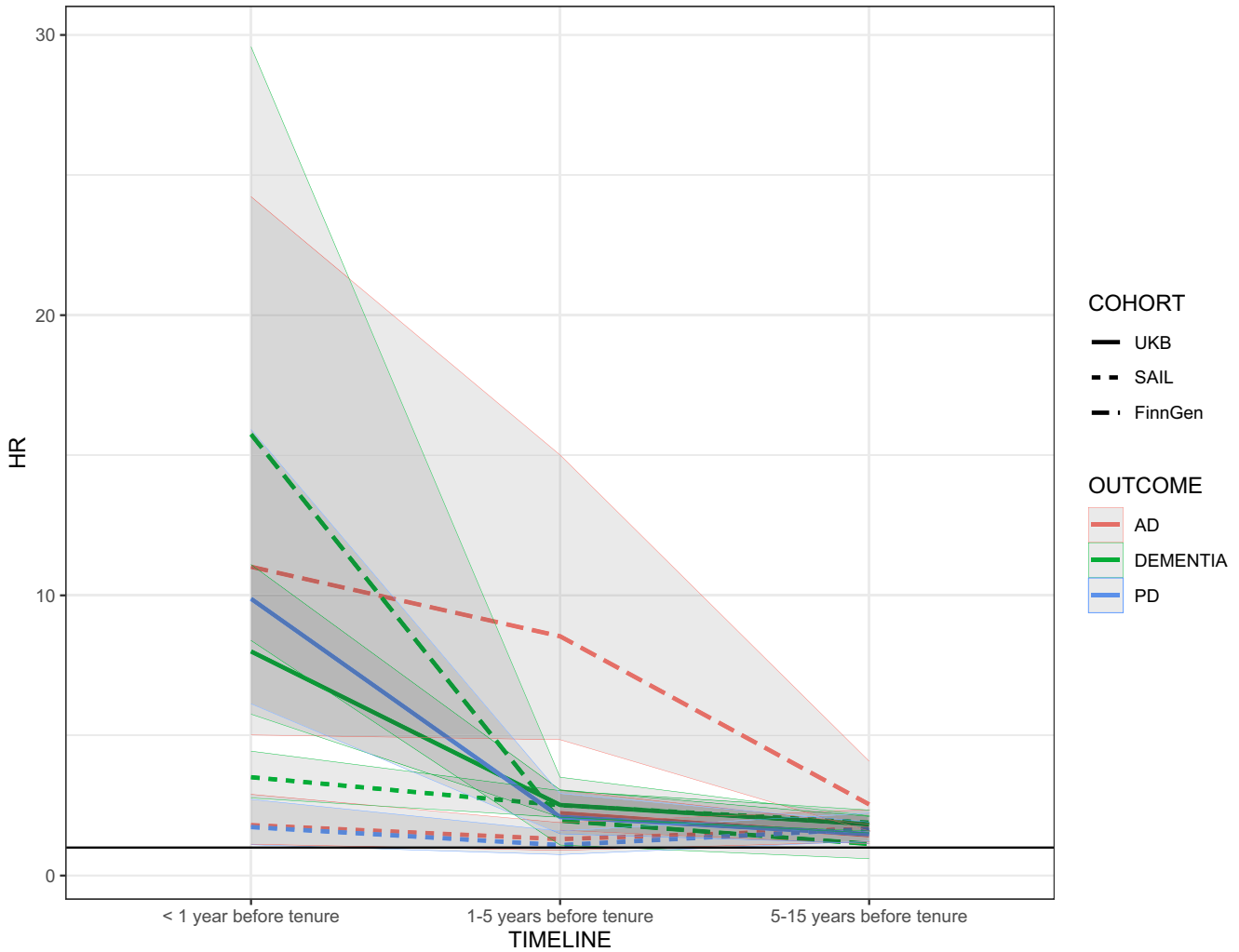


**Fig. 2 | HRs for association between sleep disorders (G47) and NDDs, split by the time between sleep disorder and NDD.** Note: There is no value for AD in UKB at <1 year before tenure, because there were only three overlapping AD/G47 and we required at least five overlapping cases in our analysis.

we were to include individuals using a CPAP machine as a sleep apnea proxy, but there would be a more significant *p*-value due to the increased numbers.

The situation was similar in UKB. We used the general practitioner (GP) clinical events records to look for CPAP encodings (Supplementary Table 1, Table 5). While ICD10 codes are available for almost the entire UKB cohort, GP clinical events data is only available for ~45% of the cohort. In addition, the GP data ends in 2017, greatly limiting the amount of data available for analysis. Similar to SAIL, a large percentage of people with a CPAP encoding in their clinical record had no ICD10 diagnosis code. (~17–18%) Overall, we found 194 individuals with a CPAP prescription in the clinical event data. This seems surprisingly low, since there are ~11,000 people with an G47.3 ICD10 code for sleep apnea. There was not enough case overlap between CPAP prescription and NDD diagnosis to conduct an analysis in UKB.

### Risk of neurodegeneration remains up to 15 years before the disease onset

In order to better assess the temporal effect on risk over fifteen years of follow-up, we split our cohorts into strata, dividing the cohorts by those who received a sleep disorder code in the EHR less than one year (Lag 0–1), one to five years (Lag 1–5), or five to fifteen (Lag 5–15) years prior to NDD diagnosis. The analysis using all available data was considered Lag 0. We then re-evaluated the risk of NDDs for the six significant pairings described above. These associations are detailed in Supplementary Table 3.

The largest subset HRs were found less than one year before diagnosis, with HRs ranging from 1.73 for G47/PD in SAIL (95% CI 1.10–2.72) to 21.24 for F51/dementia in FinnGen (95% CI 9.32–48.37). In these cases, the sleep disorders may be early symptoms of neurological disease.

However, for AD, dementia, and PD, the G47 pairing remained significant 5–15 years before diagnosis, with HRs ranging from 1.44 (for AD 95% CI 1.13–1.83 in UKB) to 2.54 (for AD 95% CI 1.59–4.08 in FinnGen). For a graphical summary of these results, see Fig. 2. We then broke the groupings down further for UKB and SAIL, looking at 5–10 (Lag 5–10) and 10–15 (Lag 10–15) year intervals. For AD and dementia, the association remained significant 5–10 years before diagnosis in both cohorts. For dementia, the G47 association also remained significant 10–15 years before diagnosis in UKB, with an HR of 1.56 (95% CI 1.27–1.94). For PD, the G47 association remained significant 5–10 years before diagnosis in UKB and 10–15 years before diagnosis in SAIL, with an HR of 1.70 and 1.84 respectively (95% CI 1.22–2.38 and CI 1.20–2.82). The G47.3 (sleep apnea) coding also remained significant for dementia and vascular dementia at 5–10 years before NDD diagnosis, with HRs ranging from 1.71 (dementia 95% CI 1.33–2.20 in UKB) to 3.25 (vascular dementia 95% CI 1.98–5.33 in SAIL).

The UKB also contains participants from England, Scotland, and Wales. While the records have been anonymized, we suspect there is some overlap with individuals in SAIL. To account for any potential overlap, we reran all the Cox regressions with the Welsh participants removed. These results are reported in Supplementary Table 2 and 3. Results were similar at

**Table 2 | Multi-exposure for G47 and F51 sleep disorders**

| ICD10 | PRIOR | OUTCOME | COHORT | Nexp vs None | HR | ci_min | ci_max | P_VAL | N_pairs | N | P_CORR | SIGNIFICANT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| G47 | Other sleep disorders | DEMENTIA | SAIL | Ordinal | 1.42 | 1.34 | 1.49 | 8.10E-37 | 311 | 6820 | 8.91E-36 | TRUE |
| G47 | Other sleep disorders | AD | SAIL | Ordinal | 1.1 | 0.99 | 1.23 | 8.06E-02 | 95 | 6820 | 1.48E-01 | FALSE |
| G47 | Other sleep disorders | PD | SAIL | Ordinal | 1.04 | 0.94 | 1.15 | 4.92E-01 | 108 | 6820 | 6.01E-01 | FALSE |
| G47 | Other sleep disorders | VASCULAR | SAIL | Ordinal | 1.56 | 1.41 | 1.73 | 2.73E-18 | 76 | 6820 | 1.50E-17 | TRUE |
| F51 | Nonorganic sleep disorders | DEMENTIA | SAIL | Ordinal | 2.19 | 1.51 | 3.18 | 3.66E-05 | 11 | 78 | 1.01E-04 | TRUE |
| G47 | Other sleep disorders | DEMENTIA | SAIL | 1 | 2.09 | 1.78 | 2.45 | 2.29E-19 | 154 | 3389 | 7.56E-18 | TRUE |
| G47 | Other sleep disorders | DEMENTIA | SAIL | 2 | 2.51 | 1.95 | 3.24 | 1.28E-12 | 60 | 1249 | 1.06E-11 | TRUE |
| G47 | Other sleep disorders | DEMENTIA | SAIL | 3+ | 2.29 | 1.87 | 2.80 | 7.71E-16 | 97 | 2182 | 8.48E-15 | TRUE |
| G47 | Other sleep disorders | VASCULAR | SAIL | 1 | 1.75 | 1.20 | 2.54 | 3.50E-03 | 28 | 3389 | 1.16E-02 | TRUE |
| G47 | Other sleep disorders | VASCULAR | SAIL | 2 | 2.02 | 1.11 | 3.65 | 2.06E-02 | 11 | 1249 | 5.65E-02 | FALSE |
| G47 | Other sleep disorders | VASCULAR | SAIL | 3+ | 3.87 | 2.78 | 5.37 | 7.13E-16 | 37 | 2182 | 8.48E-15 | TRUE |
| G47 | Other sleep disorders | AD | SAIL | 1 | 1.70 | 1.32 | 2.18 | 4.52E-05 | 61 | 3389 | 1.86E-04 | TRUE |
| G47 | Other sleep disorders | AD | SAIL | 2 | 1.09 | 0.63 | 1.89 | 7.47E-01 | 13 | 1249 | 9.91E-01 | FALSE |
| G47 | Other sleep disorders | AD | SAIL | 3+ | 1.04 | 0.67 | 1.59 | 8.69E-01 | 21 | 2182 | 9.91E-01 | FALSE |
| G47 | Other sleep disorders | PD | SAIL | 1 | 1.75 | 1.38 | 2.20 | 2.65E-06 | 73 | 3389 | 1.25E-05 | TRUE |
| G47 | Other sleep disorders | PD | SAIL | 2 | 0.97 | 0.58 | 1.65 | 9.20E-01 | 14 | 1249 | 9.91E-01 | FALSE |
| G47 | Other sleep disorders | PD | SAIL | 3+ | 0.79 | 0.52 | 1.22 | 2.87E-01 | 21 | 2182 | 5.58E-01 | FALSE |

FDR $p < 0.05$ after excluding 21,182 Welsh UKB participants; directionality and significance remained the same.

### More severe sleep disturbances impart a higher risk

In SAIL, data was available for G47-coded sleep disorders with more detail, so we used the number of times an individual received a sleep disorder ICD10 code in their EHR as a proxy for disease severity. We compared individuals risk estimates at one, two, and three or more recorded sleep disorder codes during follow-up. Those who experienced zero incidents of sleep disturbances were used as the reference group. In general, the trend for increasing risk with recurrent sleep disturbances codes was significant for dementia (HR 1.42, 95% CI 1.34–1.49, *p*-value 8.9E-36) and vascular dementia (HR 1.56, 95% CI 1.41–1.73, *p*-value 1.50E-17). This trend was mirrored for F51-coded sleep disorders in dementia as well (HR 2.19, 95% CI 1.51–3.18, *p*-value 1.01E-04). The highest trend for increasing risk with recurrent sleep disturbances was with vascular dementia, this may be due to cardiovascular comorbidities, poor diet, and lack of exercise which will also impact sleep disorders and risk of vascular dementia. The effect is also high for dementia, but this will encompass individuals with vascular dementia and is likely driven by these individuals. These association tests are detailed in Table 2.

We did not have data available to replicate this analysis in UKB or FinnGen.

### Sleep disturbances increase overall risk in individuals with low genetic risk

We also evaluated potential genetic interactions between the PRS for PD and AD with sleep disorders. Despite the fact that RBD and PD share some genetic risk factors, like SNCA and GBA, distributions of PD PRS were significantly different between cases of PD with a sleep disorder versus cases that did not have either F51 (T-test: statistic −4.16, *p*-value 3.29E-05) or G47 (T-test: statistic −3.74, *p*-value 1.89E-04) sleep disorders. In other words, if a patient had a sleep disorder and PD, they tended to have less genetic risk for PD; and if a patient did not have a sleep disorder, they tended to have more genetic risk for PD.

For AD, distributions of AD PRS (including *APOE*) were also significantly different for G47 (T-test: statistic −2.38, *p*-value 1.75E-02). If a patient had AD and a sleep disorder, they tended to have less genetic risk (lower PRS) than those without a sleep disorder. When we excluded *APOE* from the AD PRS, results were no longer significant (T-test: statistic −1.00, *p*-value 0.32). There were no significant differences in distribution for AD and F51. These results are graphically summarized in Fig. 3.

### Interactions between sleep disorders and genetic risk

The differences in genetic risk between people with and without sleep disorders led us to formally test for interactions between the PRS for each NDD and each of the significant sleep disorders, as detailed in Table 3. For AD, the sleep code G47 remained significant with the inclusion of AD PRS (with or without *APOE*) as a covariate; the interaction term was not significant in either model.

For PD, the G47 sleep code remained significant with the inclusion of PD PRS, however the interaction term was also significant. Given the known risk loci for RBD near known PD genes, it seems reasonable that there would be some interaction between the sleep disorder and the genetic risk. For the F51 model, including the PD PRS, caused the F51 code to no longer be significant; in this model the interaction term was also significant.

In both models, the interaction term's OR was less than one, however the coefficient of an interaction in a logistic model cannot be interpreted directly, because it presents the ratio of terms at different states. For example in Table 3 for the G47/PD interaction:

NDD ~ Beta(G47) + Beta(PRS) + Beta(interaction term (G47*PRS))

When G47 is 0 (no sleep disorder), the OR is the OR for the PRS.

$OR_{G47=0}$ = exp(Beta(PRS)) = exp(0.64) = 1.90

When G47 is 1 (patient has a sleep disorder), the OR is exp(Beta(PRS) + Beta(interaction term))

$OR_{G47=1}$ = exp(Beta(PRS) + Beta(interaction term)) = exp(0.64–0.32) = 1.38

With the sleep disorder, the OR is less than the term on its own. And if you take the ratio of these two odds ratios, you get the OR of the interaction term: 1.38/1.90 = 0.73

We can do the same calculations for F51:

$OR_{F51=0}$ = exp(Beta(PRS)) = exp(1.78) = 5.91

$OR_{F51=1}$ = exp(Beta(PRS) + Beta(interaction term)) = exp(1.78–1.46) = 1.38

$OR_{interaction\ term}$ = 1.38/5.91 = 0.23

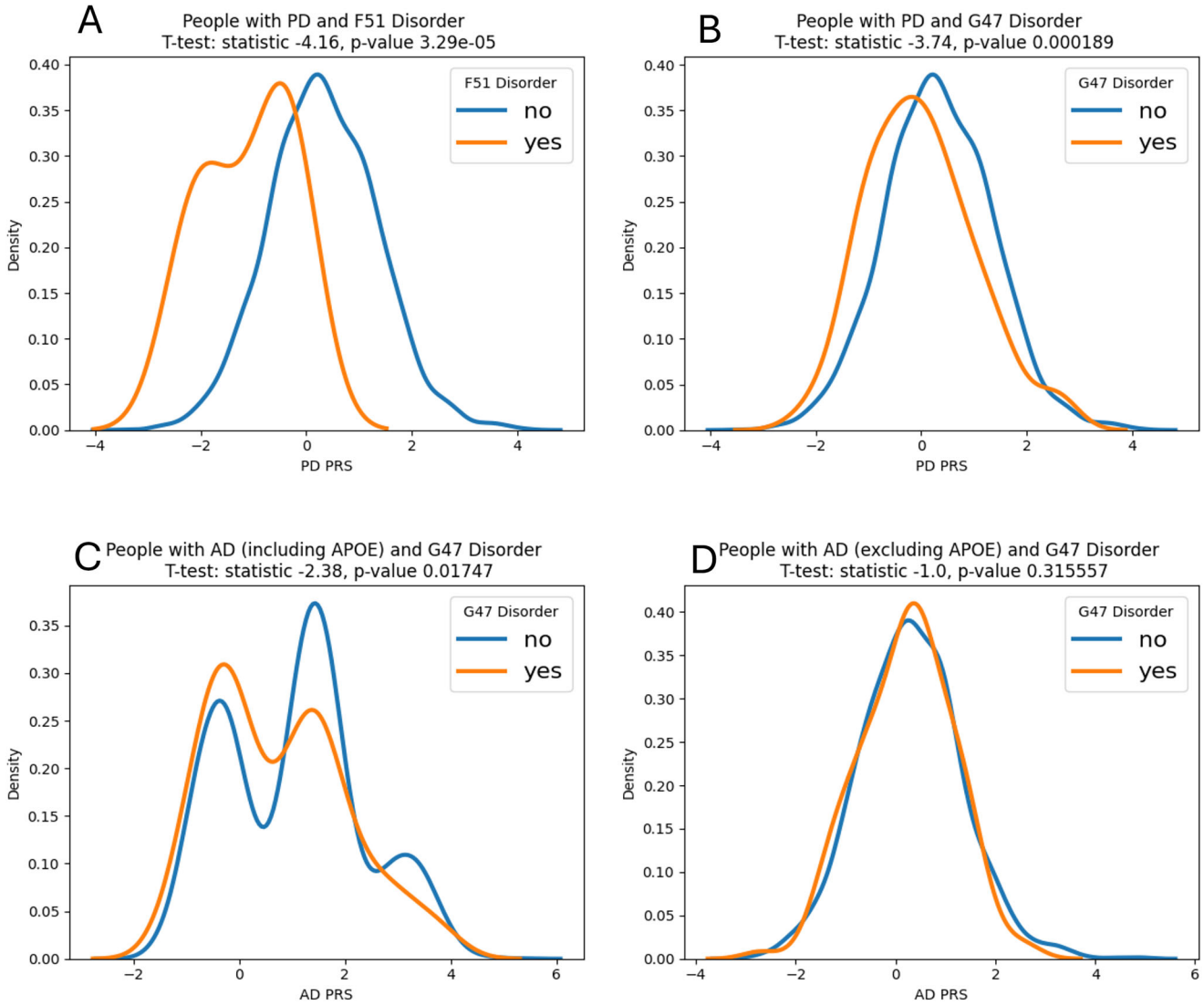


**Fig. 3 | PRS density plots for NDD cases with and without a sleep disorder. A** years before diagnosis in SAIL PD and F51; **B** PD and G47; **C** AD (including APOE) and G47; **D** AD (excluding APOE) and G47. We looked at PD PRS and AD PRS (with and without APOE) and found that PRS are differently distributed between cases of NDDs with and without sleep disturbances. These differences were significant for PD & F51, PD and G47, and AD (with APOE) and G47. AD PRS without APOE was not significant.

**Table 3 | PRS and Interactions with significant sleep disorders**

| Model | Parameter | OR | Beta | SE | 95% CI low | 95% CI high | *P*-value | Interaction significant |
|---|---|---|---|---|---|---|---|---|
| **G47 and AD PRS interaction (with APOE4)** | *G47* | 1.52 | 0.42 | 0.10 | 0.21 | 0.62 | 6.27E-05 | |
| | *interaction term* | 0.89 | −0.11 | 0.07 | −0.26 | 0.04 | 1.38E-01 | False |
| | AD PRS (Kunkle et al.) | 2.42 | 0.88 | 0.08 | 0.73 | 1.04 | 2.79E-29 | |
| **G47 and AD PRS interaction (excluding APOE4)** | *G47* | 1.34 | 0.29 | 0.09 | 0.12 | 0.47 | 1.25E-03 | |
| | *interaction term* | 0.93 | −0.07 | 0.09 | −0.25 | 0.10 | 4.25E-01 | False |
| | AD PRS (Kunkle et al.) | 1.46 | 0.38 | 0.09 | 0.20 | 0.56 | 5.07E-05 | |
| **G47 and PD PRS interaction** | *G47* | 1.50 | 0.40 | 0.09 | 0.22 | 0.58 | 1.06E-05 | |
| | *interaction term* | 0.73 | −0.32 | 0.09 | −0.50 | −0.14 | 5.76E-04 | True |
| | PD PRS (Nalls et al.) | 1.90 | 0.64 | 0.10 | 0.45 | 0.83 | 4.54E-11 | |
| **F51 and PD PRS interaction** | *F51* | 2.38 | 0.87 | 0.56 | -0.22 | 1.96 | 1.19E-01 | |
| | *interaction term* | 0.23 | −1.46 | 0.42 | −2.29 | −0.64 | 5.24E-04 | True |
| | PD PRS (Nalls et al.) | 5.91 | 1.78 | 0.42 | 0.95 | 2.60 | 2.60E-05 | |

So the fact that the interaction term is <1 is not saying that variable is protective; instead, it is a representation that in our data we typically see a lower PRS in the presence of a sleep disorder.

We also tested each of the interaction models with age at the onset of either PD or AD as the outcome and no interaction terms were significant[21].

## Discussion

In this work, we have gained insights into the complex relationship between sleep disturbances and late-life risk of neurodegeneration. While previous studies have investigated the associations between sleep and neurodegenerative diseases, this is the first large-scale survey that has shown replicated associations for multiple NDDs across multiple biobanks. This risk remained significant for up to 10–15 years before NDD diagnosis for dementia, and 5–10 years for AD and PD. This is an important finding, since successful diagnosis and treatment of sleep disorders may be used as an early intervention to reduce an individual's risk of NDDs. The severity of sleep disorders tended to increase risk as well. Finally, we looked into the interplay between sleep-related risk and genetically derived risk, providing evidence that sleep disorders and genetics are likely separate mechanisms leading to the same outcome in AD, but seem to interact in PD to some degree.

Maintaining proper sleep hygiene can potentially help reduce an individual's risk of developing a neurodegenerative disease with a number of lifestyle, chemical, and mechanical interventions that are widely available (i.e. medications, CPAP machines). The circadian clock and sleep can influence several key processes involved in neurodegeneration, suggesting that these systems might be manipulated to promote healthy brain aging[9]. In addition, there is some evidence that positive airway pressure therapy for sleep apnea is associated with a lowered risk of AD or dementia diagnosis, in a retrospective study of individuals aged 65+[22].

Medications used in neurodegenerative conditions can also affect sleep architecture. Dopaminergic medications used in PD can improve motor symptoms, but may lead to sleep disturbances, including insomnia and restless leg syndrome. A review of recent publications highlights that some antidepressant medications can lead to insomnia[23]. A review and meta-analysis of studies investigating the impact of benzodiazepines on dementia risk found that dementia risk is increased with benzodiazepine use[24]. Cholinesterase inhibitors, such as donepezil, can improve cognitive and behavioral symptoms in people with Alzheimer's but may have side effects like vivid dreams, nightmares, and insomnia, as seen from patient charts at the Tallahassee Memorial Healthcare Neuroscience Center[25]. However, medications used to improve sleep may slow or stop the progression of neurodegeneration[26,27].

Sleep disorders were still mostly significantly associated with the tested NDD phenotype even after correction for PRS, suggesting that while there are some known common genetic risk factors for sleep disorders and disease (such as *GBA1* shared between RBD and PD)[28], the risk observed between sleep disorders and NDDs does not appear to be due to underlying genetic risk factors alone. Not only does this report show a propensity for sleep hygiene to help influence improved cognitive outcomes in later life, but our analysis suggests interesting ramifications for precision medicine research. For example, if designing a clinical trial to look at the use of CPAP in NDD prevention, our results suggest it would make sense to stratify patients into low and high-genetic-risk groups; it is possible that treatment with CPAP would be protective in those with low genetic PRS, while having no effect in the high genetic risk group[29].

This report is not without its limitations. Summary data level access to FinnGen is excellent, although participant-level access could have facilitated additional modeling efforts. All of our datasets use diagnoses based on medical billing codes and not bloodwork or other assays. This coupled with data sparsity, can cause issues with some analyses. For example, the ICD10 code for RBD is G47.52, which means it should be included in the G47 coding. However, while we were able to break out the specific coding for sleep apnea (G47.3), the specific code for RBD does not appear in UKB, so we were unable to analyze it separately. It is likely that RBD in UKB would be included in G47.8 "Other sleep disorders", which does show a highly significant association with PD (HR 8.07, 95% CI 4.03–16.17, corrected *p*-value 1.6e-08).

When comparing the results of SAIL and UKB, the differences between the two datasets in the results may be due to available diagnosis codes and differences in sample makeup, such as age distribution and participant selection bias. In particular, incidence of dementia in the UKB may be different compared to the SAIL dataset because the UKB participants are volunteers who have not yet reached old age and are self-selected. We also did not account for the duration and for medication usage when considering severity, using only history and numbers of sleep disorders as a proxy for severity.

In addition, all the cohorts analyzed are predominantly European ancestry and may not be globally representative. Finally, we only had access to genetic data in one dataset on a participant level. Therefore, tests of interaction and independence of genetics and sleep related risk have yet to be formally replicated. There may be some differences between the three cohorts due to regional differences, although this is likely to be minor. All cohorts are European: SAIL is from Wales, UK, UKB is from the UK and incorporates individuals in Wales, and FinnGen is the Finnish population. Regional differences are most likely to occur between FinnGen and SAIL/UKB due to different healthcare systems and therefore data collection. There may also be small differences caused by the fact that individuals are recruited to UKB and FinnGen, whereas SAIL is population-based and therefore representative of the general population.

Finally, sleep disorders are often underdiagnosed, so we may have individuals in our 'non-exposed' group who have experienced a sleep disorder. Therefore, the results presented may be diluted and the true impact of sleep disorders may be greater. The high impact of sleep apnea seen here is likely due to sleep disorder diagnoses being from hospital records, which will better capture specific sleep disorders. In SAIL, we additionally investigated diagnoses based on GP records, but the results remained consistent.

The lack of clean data related to CPAP in the three biobanks we examined, e.g. small numbers of people recorded as having both sleep apnea and using a CPAP machine, suggests this as an area for future research by further understanding the data collection process.

Medications used to treat sleep disorders by improving a person's sleep may slow or prevent neurodegeneration progression, particularly with the finding that risk is maintained >5 years prior to NDD diagnosis, therefore it would be interesting to incorporate these medications into our models to determine their impact on neurodegeneration. The duration of treatment can also be incorporated which may be an additional measure for the severity of the sleep disorder. This study only captures data over 20 years, but with additional data, we could investigate whether the impact of sleep disorders on NDDs is sustained for >10/20 years, as this would capture the period prior to brain changes in dementia development.

Future studies may utilize data from wearable technology which may mitigate the issue of sleep disorder underdiagnosis, although this data would need to be matched to medical records to ensure availability of an individual's neurodegenerative disorder status.

The cohorts in this study are predominantly of European ancestry, future studies would benefit from large electronic health and genetic data being available in diverse ancestry groups.

## Methods

This study uses data from three biobanks: the Secure Anonymised Information Linkage (SAIL) databank[30], as well as the UK Biobank (UKB)[31] and FinnGen datasets[32]. SAIL is a repository of medical health records from Wales, UK, covering ~80% of all individuals living in Wales between 1970 and 2019 (Application Number 0998). UKB data were applied for and accessed via the UKB research and analysis portal (Application Number 33601), which hosts genotyping data from nearly 500,000 individuals from the UK. For controls in these cohorts, we used a subset of older (baseline age greater than 45 years) unrelated individuals of European ancestry who did not have an NDD diagnosis or a family history of NDDs. We selected older individuals in the control group to ensure they had a chance of developing

dementia during the course of the study. Summary statistics from FinnGen, a nationwide Finnish biobank with genotyping data available for over 377,000 individuals, including hazard ratios for sleep disorder endpoints, were downloaded through FinnGen's Ristey's portal in May 2023, to further replicate our results.

NDD diagnosis was determined by ICD10 codes; in UKB, for example, their algorithmically defined Alzheimer disease cohort includes those with either F00 (dementia in Alzheimer's disease) or G30 (Alzheimer's disease) in their medical records. The first reported date was used as the date of diagnosis. We used these ICD10 codes to create similar cohorts in SAIL and UKB. See Supplementary Table 1, Tab 1, for a listing of all codes used in NDD cohort creation.

We used the first reported date of an ICD10 sleep code as a sleep disorder exposure, excluding any ICD10 codes occurring after an NDD diagnosis. These were grouped into two categories, F51, which contained sleep disorders not caused by a known condition (such as F51.4 sleep terrors) and G47, which is a broader category that contains sleep disorders that may have an underlying medical cause (such as G47.3 sleep apnea). In the ICD10 coding system, F51 falls under mental health disorders, while G47 falls under diseases of the nervous system. A complete listing of all codes we included in our analysis can be found in Supplementary Table 1, Tables 2–4. Any NDD/sleep disorder association with less than five pairs (i.e. individuals with a sleep disorder diagnosis who also later developed an NDD) were excluded to reduce low statistical power analyses.

All statistical analyses, including logistic regressions, cox proportional hazard models, and polygenic risk score associations were adjusted for age, sex, and socioeconomic status when possible. These covariates were selected to ensure consistency between results obtained from FinnGen, UKB, and SAIL. All reported p-values were adjusted using false discovery rate correction applied to the tests within each cohort separately.

A summary of the study design and analyses can be found in Fig. 4. See Table 4 for case numbers and sex breakdowns for each of the six NDDs. The

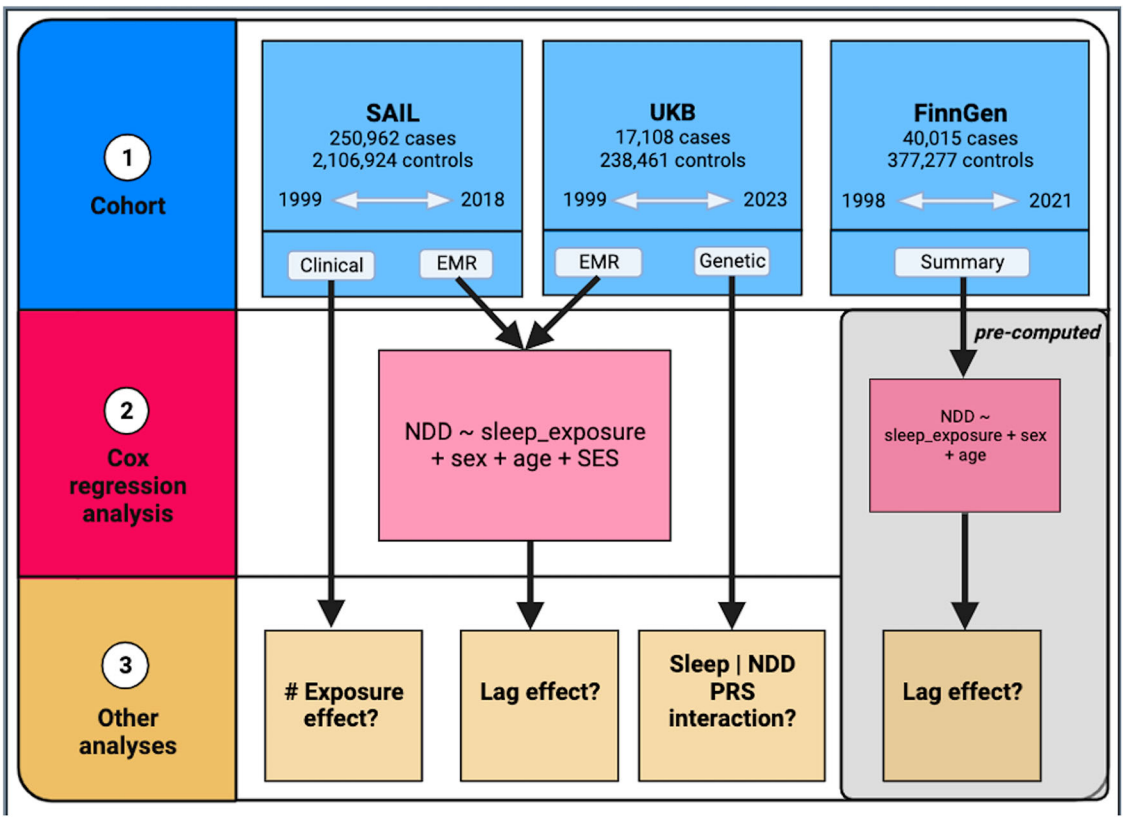


**Fig. 4 | The study design of this project.** The top panel presents the cohort information; the middle panel describes primary analyses; the bottom panel details secondary analyses.

**Table 4 | Summary of participants in the study**

| | FinnGen | | | SAIL | | | UKB | | |
|---|---|---|---|---|---|---|---|---|---|
| NDD | All | Female | Male | All | Female | Male | All | Female | Male |
| AD | 10520 | 4620 | 5900 | 38,973 | 24,743 | 14,230 | 3592 | 1880 | 1712 |
| ALS | 531 | 204 | 327 | NA | NA | NA | 556 | 245 | 311 |
| dementia | 19157 | 8220 | 10,937 | 145,081 | 86,320 | 58,761 | 7835 | 3744 | 4091 |
| MS | 2409 | 1819 | 590 | 8240 | 5472 | 2768 | 515 | 351 | 164 |
| PD | 4681 | 1858 | 2823 | 24,270 | 11,387 | 12,883 | 2891 | 1060 | 1831 |
| vascular | 2717 | 907 | 1810 | 34,398 | 19,980 | 14,418 | 1,719 | 714 | 1005 |
| controls[a] | 377,277 | 210,870 | 166,407 | 2,106,924 | 1,129,391 | 977,533 | 238,461 | 126,898 | 111,563 |

[a]In FinnGen, the control numbers listed below are the general population count (total); in the UKB these include exclusions described in the methods section such as concurrent neurodegenerative disease diagnoses.

mean age of participants in FinnGen was 53 years and the mean age at baseline of participants in our subset of SAIL and UKB individuals was 61 and 58 years, respectively.

Our initial study design included a discovery and replication phase, using time-aware data from three biobank scale datasets: SAIL, UKB, and FinnGen. Details of all codings used are given in Supplementary Table 1.

### SAIL databank

The Secure Anonymised Information Linkage (SAIL) databank is a virtual platform that facilitates the linkage of various medical datasets using anonymised IDs. Diagnoses of disorders in SAIL came from the Patient Episode Database for Wales (PEDW), records from clinicians and hospital staff, and from the Welsh Longitudinal General Practitioner dataset (WLGP). Demographic data on sex, age, address, and death were sourced from the Welsh Demographic Services Database (WDSD), and from WLGP.

To identify disorder diagnoses, PEDW was queried using ICD10 codes, and WLGP using NHS read codes (CVT2, 3). We did not include neurodegenerative disorders from the outpatient data (OPDW) since only 0.1% of dementia cases are present in this data, and it only provides the attendance date which would not be a robust diagnosis date. Due to the lack of accurate diagnosis date, we also did not include dementia diagnoses from death data (ADDES).

The study period used was January 1st 1999 to December 31st 2018. Therefore, for inclusion, individuals needed to be alive at the start of 1999. Additionally, to ensure that all included individuals had a chance of developing dementia before the end of the study period, we decided to include only those who were aged at least 45 on January 1st 1999. We also excluded any individuals missing age and sex information and individuals who did not have a Welsh address.

### UK Biobank

The UK Biobank (UKB) is a large prospective cohort of individuals from the UK recruited between 2006 and 2010 [20]. Diagnoses of disorders in UKB were derived using ICD10 codes as in SAIL. Information on sex and Townsend deprivation scores was available for all individuals. We left-censored any data occurring before January 1, 1999, to match SAIL's records and right-censored any data occurring after NDD diagnosis. Any ICD10 codes with fewer than five cases were excluded from analysis.

Individuals from the UK Biobank were selected for this study if they self-reported as white British and were of similar genetic ancestry by principal component analysis (UK Biobank field 22006) and were unrelated (kinship coefficient <0.0884). Genetic data in UKBB is already imputed using the haplotype reference consortium (HRC)[21].

### FinnGen

We downloaded time-aware (stemming from longitudinal cohort data) risk estimates from the FinnGen Risteys R10 portal for six neurodegenerative diseases: Alzheimer's disease, amyotrophic lateral sclerosis, generalized dementia, multiple sclerosis, Parkinson's disease, and vascular dementia. FinnGen endpoints are collections of one or more ICD10 codes, as defined by FinnGen clinical expert groups. In FinnGen all concurrent endpoints include at least ten cases.

### Regression models

The Cox proportional hazard model was used to calculate HRs for Tables 1 and 2, Supplementary Tables 2 and 3, and Figs. 1 and 4. The time scale used in the Cox proportional hazards model was an individual's age (age at onset, age at death, or age at the end of the study, whichever occurred first) and the event was the presence/absence of an NDD. Statsmodels GLM was used for Table 3. All code can be found at: [https://github.com/NIH-CARD/NDD_x_sleep]. In Table 1, the HRs from the Cox proportional hazards model were meta-analyzed together using the inverse variance method and random effects model. The individual HRs per cohort can be found in Supplementary Table 2.

### Polygenic risk score calculation

Polygenic risk scores (PRSs) were calculated using summary statistics for AD (Kunkle et al.) and PD (Nalls et al.)[33,34]. Summary statistics did not contain UK Biobank individuals. PRSs were adjusted for principal components to allow for population stratification. Interaction terms were defined as a binarized G47 or F51 code with one added to it to avoid scaling issues when multiplied by the PD or AD PRS (the PD and AD PRS were Z-scaled with no true zero values at floating point limits).

## Data availability

This paper analyzes existing, publicly available data. FinnGen data is available in the link here [https://risteys.finngen.fi/] and was last accessed in July 2023, using Risteys R10. The UKB data is available here [https://www.ukbiobank.ac.uk/] and last accessed on August 30th 2023. Both the discovery and replication phase dataset are covered by the relevant national IRB groups. Data and Code Availability: Notebooks containing code used in this analysis can be found in the GitHub link here: [https://github.com/NIH-CARD/NDD_x_sleep]. This paper analyzes existing, publicly available data. In addition, complete summary statistics describing these data/processed datasets derived from these data have been deposited in the supplementary materials connected to this publication and are also available on figshare (https://doi.org/10.6084/m9.figshare.28787675) and are publicly available as of the date of publication.

Received: 23 September 2024; Accepted: 31 March 2025;
Published online: 27 May 2025

## References

1. World Health Organization, *Guidelines on Physical Activity, Sedentary Behaviour and Sleep for Children Under 5 Years of Age* (World Health Organization, 2019).
2. Baglioni C., Espie C. A., Spiegelhalder K., Gavriloff D., Riemann D. Recommendation of the European Academy for Cognitive-Behavioural Therapy for Insomnia (CBT-I) for high quality training for health professionals. *Cognitive-Behavioural Therapy For Insomnia (CBT-I) Across The Life Span*. 243–250 (John Wiley & Sons Ltd, 2022).
3. 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).
4. Lananna, B. V. & Musiek, E. S. The wrinkling of time: Aging, inflammation, oxidative stress, and the circadian clock in neurodegeneration. *Neurobiol. Dis* **139**, 104832 (2020).
5. Emamian, F. et al. The association between obstructive sleep apnea and Alzheimer's Disease: a meta-analysis perspective. *Front. Aging Neurosci.* **8**, 78 (2016).
6. Hung, C. M. et al. Risk of dementia in patients with primary insomnia: a nationwide population-based case-control study. *BMC Psychiatry* **18**, 38 (2018).
7. Potvin, O. et al. Sleep quality and 1-year incident cognitive impairment in community-dwelling older adults. *Sleep* **35**, 491–499 (2012).
8. Ju, Y. E., Lucey, B. P. & Holtzman, D. M. Sleep and Alzheimer disease pathology-a bidirectional relationship. *Nat. Rev. Neurol.* **10**, 115–119 (2014).
9. Musiek, E. S. & Holtzman, D. M. Mechanisms linking circadian clocks, sleep, and neurodegeneration. *Science* **354**, 1004–1008 (2016).
10. Rigat, L., Ouk, K., Kramer, A. & Priller, J. Dysfunction of circadian and sleep rhythms in the early stages of Alzheimer's disease. *Acta Physiol.* **238**, e13970 (2023).
11. Ooms, S. & Ju, Y. E. Treatment of sleep disorders in dementia. *Curr. Treat Options Neurol.* **18**, 40 (2016).
12. McCurry, S. M. et al. Characteristics of sleep disturbance in community-dwelling Alzheimer's disease patients. *J. Geriatr. Psychiatry Neurol.* **12**, 53–59 (1999).
13. Spira, A. P. et al. Self-reported sleep and beta-amyloid deposition in community-dwelling older adults. *JAMA Neurol.* **70**, 1537–1543 (2013).

14. Tandberg, E., Larsen, J. P. & Karlsen, K. A community-based study of sleep disorders in patients with Parkinson's disease. *Mov. Disord.* **13**, 895–899 (1998).
15. Postuma, R. B. et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. *Brain* **142**, 744–759 (2019).
16. Krohn, L. et al. Genome-wide association study of REM sleep behavior disorder identifies polygenic risk and brain expression effects. *Nat. Commun.* **13**, 7496 (2022).
17. Dodet, P. et al. Sleep disorders in Parkinson's disease, an early and multiple problem. *NPJ Parkinsons Dis.* **10**, 46 (2024).
18. Elwood, P. C. et al. Sleep disturbance and daytime sleepiness predict vascular dementia. *J. Epidemiol. Community Health* **65**, 820–824 (2011).
19. Carnicka, Z. et al. Sleep disorders in patients with multiple sclerosis. *J. Clin. Sleep Med.* **11**, 553–557 (2015).
20. Boentert, M. Sleep disturbances in patients with amyotrophic lateral sclerosis: current perspectives. *Nat. Sci. Sleep* **11**, 97–111 (2019).
21. VanderWeele, T. J. & Knol, M. J. A tutorial on interaction. *Epidemiol. Methods* **3**, 33–72 (2014).
22. Dunietz G. L., Chervin R. D., Burke J. F., Conceicao A. S., Braley T. J. Obstructive sleep apnea treatment and dementia risk in older adults. *Sleep*. **44**, zsab076 (2021).
23. Wichniak, A., Wierzbicka, A., Walecka, M. & Jernajczyk, W. Effects of Antidepressants on Sleep. *Curr Psychiatry Rep.* **19**, 63 (2017).
24. Ferreira, P., Ferreira, A. R., Barreto, B. & Fernandes, L. Is there a link between the use of benzodiazepines and related drugs and dementia? A systematic review of reviews. *Eur Geriatr. Med.* **13**, 19–32 (2022).
25. Agboton, C. et al. Impact of nighttime donepezil administration on sleep in the older adult population: a retrospective study. *Mental Health Clin.* **4**, 257–259 (2014).
26. Shen, Y. et al. Circadian disruption and sleep disorders in neurodegeneration. *Transl. Neurodegener.* **12**, 8 (2023).
27. Xiong, X. et al. Research advances in the study of sleep disorders, circadian rhythm disturbances and Alzheimer's disease. *Front. Aging Neurosci.* **14**, 944283 (2022).
28. Krohn, L. et al. GBA variants in REM sleep behavior disorder: a multicenter study. *Neurology* **95**, e1008–e1016 (2020).
29. Minikel, E. V. et al. Refining the impact of genetic evidence on clinical success. *Nature* **629**, 624–629 (2024).
30. Jones, K. H., Ford, D. V., Thompson, S. & Lyons, R. A. A profile of the SAIL databank on the UK secure research platform. *Int. J. Popul. Data Sci.* **4**, 1134 (2019).
31. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
32. Kurki, M. I. et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* **613**, 508–518 (2023).
33. Kunkle, B. W. et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).
34. Nalls, M. A. et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* **18**, 1091–1102 (2019).


## Acknowledgements

This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging (NIA), National Institutes of Health, Department of Health and Human Services; project number ZO1 AG000535, as well as the National Institute of Neurological Disorders and Stroke. This work was also supported by the Dementia Research Institute [UKDRI supported by the Medical Research Council (UKDRI-3003), Alzheimer's Research UK, and Alzheimer's Society], Welsh Government, Joint Programming for Neurodegeneration (MRC: MR/T04604X/1), Dementia Platforms UK (MRC: MR/L023784/2) and MRC Centre for Neuropsychiatric Genetics and Genomics (MR/L010305/1). This research has been conducted using the UK Biobank Resource under application number 33601. We want to acknowledge the participants and investigators of the FinnGen study and thank them for their hard work and generosity. This study makes use of anonymized data held in the Secure Anonymised Information Linkage (SAIL) Databank. This work uses data provided by patients and collected by the NHS as part of their care and support. We would also like to acknowledge all data providers who make anonymized data available for research. We wish to acknowledge the collaborative partnership that enabled the acquisition and access to the de-identified data, which led to this output. The collaboration was led by the Swansea University Health Data Research UK team under the direction of the Welsh Government Technical Advisory Cell (TAC) and includes the following groups and organizations: the SAIL Databank, Administrative Data Research (ADR) Wales, Digital Health and Care Wales (DHCW), Public Health Wales, NHS Shared Services Partnership (NWSSP) and the Welsh Ambulance Service Trust (WAST). All research conducted has been completed under the permission and approval of the SAIL independent Information Governance Review Panel (IGRP) project number 0998.

## Author contributions

E.S., K.S.L., H.L.L., V.E.P., M.A.N., and H.I. designed the study; E.S., K.S.L., H.L.L., C.B., V.E.P. and M.A.N. acquired and processed data; E.S., K.S.L., M.A.N., H.I., V.E.P. and H.L.L. analyzed data; all authors (E.S., K.S.L, J.H., H.I., M.J.K., N.K., F.F., C.W.S., A.S., L.J., S.B.C., C.B., A.S., V.E.P., H.L.L., M.A.N.) contributed to interpreting data and writing and editing of the manuscript.

## Competing interests

K.S.L., H.L.L., H.I., F.F., C.W.S., L.J., and M.A.N.'s participation in this project was part of a competitive contract awarded to Data Tecnica International LLC by the National Institutes of Health to support open science research. M.A.N. also currently serves on the scientific advisory board at Clover Therapeutics and is an advisor and scientific founder at Neuron23 Inc.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s44400-025-00008-0.

**Correspondence** and requests for materials should be addressed to Valentina Escott-Price or Hampton L. Leonard.