Sleep trajectories and mediators of poor sleep: Findings from the longitudinal analysis of 41,094 participants of the UK Biobank cohort

Yaqoot Fatima[1, 2*], Romola S Bucks[3], Abdullah A Mamun[1], Isabelle Skinner[2], Ivana Rosenzweig [4], Guy Leschziner[5, 6], Timothy C Skinner[2], [7, 8]

1. Institute for Social Science Research, University of Queensland, Brisbane, Australia
2. Centre for Rural and Remote Health, James Cook University, Mount Isa, Australia
3. School of Psychological Science, University of Western Australia, Perth, Australia
4. Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, United Kingdom
5. Sleep Disorders Centre, Guy's and St Thomas' NHS Foundation Trust, Guy's Hospital, London, United Kingdom
6. Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
7. University Department of Rural Health, La Trobe University, Bendigo, Australia
8. Institut for Psykologi, Center for Sundhed of Samfund, Københavns Universitet, Øster Farimagsgade, København K. Denmark

*Address for Correspondence

Dr Yaqoot Fatima

Institute for Social Science Research

University of Queensland, Indooroopilly, Queensland 4068, Brisbane, Australia

Email: fatima.yaqoot@uq.edu.au

Telephone: +61 7 344 31021

Abstract

Study Objectives: To explore sleep trajectories and identify the risk factors and mediators of poor sleep in middle-aged adults.

Methods: Group-based multi-trajectory modelling was applied to the three waves of sleep data the from UK Biobank cohort to identify latent trajectories of sleep and group characteristics. Self-reported sleep duration, sleep problems (based on insomnia symptoms, snoring and trouble waking up) and daytime sleepiness (based on daytime tiredness and sleepiness) were included in the trajectory analyses. Multinomial logistic regression and mediation analysis were used to identify the main factors associated with poor sleep.

Results: Analysis of sleep data from 41,094 participants (51.9% females) with a median age of 57 years (interquartile range 50-62 years) identified three distinct trajectories of sleep: healthy sleepers (40.8%); borderline poor sleepers (31.6%); and poor sleepers (27.6%). Socio-economic disadvantage, ethnic minority background, shift work, unhealthy lifestyle, poor health, depressive symptoms and obesity were the main risk factors associated with poor sleep. Around a third of the total effect of socio-economic deprivation on poor sleep was mediated through depressive symptoms.

Conclusions: The distinct groups with differential risk for developing sleep issues and stable sleep trajectories highlight the non-transient nature of sleep issues. Early management of depressive symptoms can help in reducing the future burden of poor sleep. Due to the increased risk of poor sleep, people from socio-economically deprived groups, particularly females from ethnic minorities, should be the highest priority for interventions aiming to improve population sleep health.

Keywords: sleep trajectories, poor sleep, latent class analysis, mediation, socio-economic disadvantage, UK Biobank

Highlights

- Poor sleep is a relatively stable phenomenon; therefore, early intervention is required for sleep improvement, rather than assuming that poor sleep is a temporary problem that will remit spontaneously.
- Around 1/3rd of the study participants who reported sleeping for the recommended duration, reported considerable burden of sleep problems and daytime sleepiness.
- To optimise sleep health, we need to look beyond sleep duration and focus on the whole spectrum of sleep health issues
- In people from socio-economic deprived groups, depressive symptoms are the main mediators of poor sleep

Introduction

Epidemiological evidence from large population studies consistently highlights the rising prevalence of poor sleep and its role in the growing burden of chronic health conditions. [1,2] Many cross-sectional and longitudinal studies have endeavoured to quantify the degree to which poor sleep is a problem and identify the antecedents of poor sleep, so that appropriate prevention and management strategies can be instituted. [3,4] Unfortunately, the bulk of the evidence on the growing burden of poor sleep is predominantly solely based on measures of sleep duration. In real-world settings, poor sleep manifests in the form of concurrent indicators, e.g., short sleep, poor quality of sleep, snoring, problems in initiating and maintaining sleep, daytime sleepiness etc. [5,6] And it is the accumulation of multiple indicators, rather than any single indicator, which accounts for most of the negative health outcomes.[7] Therefore, failing to recognise the concurrence of sleep issues may impact the reliable assessment of the burden of poor sleep and associated outcomes.

Some studies have explored concurrent sleep issues, either in the form of a global score derived from validated sleep questionnaires, or adjusting for other sleep dimensions in regression models. [8] However, these studies mainly assessed changes from baseline to follow-up, using a variable-centred approach, assuming poor sleepers are a homogenous population. [9] The variable-centred approach assumes that irrespective of their characteristics, all individuals at a certain level of risk factor are at equal risk of adverse outcome. Therefore, the association between a risk factor and outcome is the same across the entire population. [10] However, recognising the individual and combinatorial impact of different individual and environmental factors in varying risk levels for developing poor sleep, it can be posited that poor sleepers are a heterogeneous population. [11,12] Further, within that population, there are inherent groups with varying degrees and nature of poor sleep, based on the intersection of various risk factors. [13] A person-centred framework overcomes the limitations of variable-

centred models by exploring development in the light of multiple person-environment interactions and identifying different hidden groups within a seemingly homogenous population. [14] Subdividing the study populations into smaller "groups" offers greater ecological validity and accounts for the concurrence and interplay among multiple risk factors in individuals. [10]

Latent class analysis (LCA), is a person-centred method for identifying subgroups in a given population by simultaneously considering multiple factors to reveal constellations in the data. [15] Longitudinal latent class analysis is an extension of LCA that captures the shape and form of the trajectories of health conditions and explains how health trajectories develop over time. [16] Using a latent class trajectory analysis helps in examining the divergent trends in specific subgroups of individuals with varying risk levels, which in turn facilitates prioritising and tailoring appropriate interventions for different groups.

One of the challenges in LCA-based modelling is the requirement of a large sample for estimating model parameters, especially when the subgroups are complex and less distinct.[17] The UK Biobank is a large population-based study offering rich longitudinal data on multiple sleep dimensions and important factors associated with poor sleep. Applying latent trajectory analysis to the UK Biobank sleep data offers a unique opportunity to address some of the limitations of existing studies on adult sleep. Hence, the primary aim of this study was to understand how co-occurring sleep issues interact and cluster together and lead to poor sleeper subgroups while exploring the trajectories of those groups over time.

In addition to identifying the phenotypes of poor sleep, it is also important to identify the factors associated with poor sleep. [18] Particularly in disadvantaged and vulnerable groups, where the disproportionately higher burden of poor sleep and associated cardio-metabolic and mental health conditions warrants urgent attention. [19,20]. Identification of key risk factors and mediators of poor sleep can facilitate the early institution of appropriate interventions and

help in reducing the burden of poor sleep and associated health outcomes. Previous research has highlighted that predisposing factors such as low socio-economic status, and poor health increases the vulnerability for poor sleep.[21] Precipitating factors such as major life events trigger the onset of sleep issues, and propagating factors such as unhealthy lifestyle support the continuity of poor sleep.[22] It is also well established that health inequity and occupational class are a function of socio-economic disadvantage. [23-25] However, it remains to be explored whether, in disadvantaged groups, the high burden of poor sleep is driven by socio-economic disparities or health issues and occupational class play a mediating role. Therefore, the second aim of this study was to identify whether poor health, obesity and shift work contribute to sleep health disparities in people from socio-economically deprived groups and identify the opportunities for sleep health improvement.

Methods

Study participants

This research was conducted using the UK Biobank data (Application Number 19705), a prospective study of 502, 618 middle-aged adults (37-73 years) aiming to improve the prevention, diagnosis, and treatment of chronic conditions. [26] Participants were recruited from 22 assessment centres in England (89%), Scotland (7%) and Wales (4%) between 2006 and 2010 to collect data on demographic and lifestyle characteristics, medical history and self-rated health. Physical measurements were also taken, and participants provided blood and urine samples. Participants living ≤35 km from the Stockport assessment centre were invited to two repeat assessments, first between December 2009 and June 2013 (n = 20,346) and then again between April 2014 and November 2016 (n = 35,540). The UK Biobank study has approval from the North West Multi-centre Research Ethics Committee, the Patient Information Advisory Group, and the Community Health Index Advisory Group. Further

details on the study design, sampling, data collection, and ethics committee approval are detailed elsewhere. [27]

Study Measures

While sleep variables from all three time points were used to identify sleep trajectories, only the baseline data for the risk factors and mediators of poor sleep were used in the analysis

Sleep variables

The data on the following six self-reported items, i.e., sleep duration; insomnia symptoms; trouble waking up; snoring; daytime sleepiness and daytime tiredness were extracted to study the trajectories of sleep. Insomnia symptoms were assessed by asking, "Do you have trouble falling asleep at night or do you wake up in the middle of the night?" with responses "never/rarely", "sometimes", and "usually". Participants' difficulty in waking up was assessed by asking, "On an average day, how easy do you find getting up in the morning" with the following response options: "not at all easy", "not very easy," "fairly easy", and "very easy". Participants were also asked, "Does your partner or a close relative or friend complain about your snoring" answering with "yes" or "no." Sleep duration was recorded by asking the following question "About how many hours sleep do you get in every 24 hours? Daytime sleepiness was assessed based on answers to the following questions: "how likely you are to doze off or fall asleep during the daytime when you don't mean to (e.g., when working, reading or driving) with responses "never/rarely", "sometimes", and "often." Daytime tiredness was assessed by asking "Over the past two weeks, how often have you felt tired or had little energy?" with the following response options: "not at all", "several days", "more than half the days", and "nearly every day."

Three broad indicators of poor sleep, i.e., sleep duration, sleep problems and daytime sleepiness were used to track sleep trajectories. Sleep duration was used as a continuous variable. In health surveys, compared with single item based assessment, constructs derived

from multiple items are reported to offer improved reliability for capturing health issues and reduced measurement error.[28] Therefore, a sleep problem construct reflecting the overall poor quality of sleep and daytime sleepiness construct reflecting insufficient sleep were used in analyses. The sleep problems construct was generated by computing the mean scores of insomnia symptoms, snoring and trouble waking up, daytime sleepiness was generated by computing mean scores of daytime tiredness and sleepiness. For all sleep items, responses "do not know" 'or "prefer not to answer" were excluded from the analysis.

Risk factors and mediators of poor sleep

Socio-demographic factors: participant's age; gender; ethnicity categorised as "white," and "ethnic minority" (Black, Asian and mixed), gross annual household income classified as "<£18 000," "£18000–30 999," "£31 000–51 000," "£52 000–100 000," and">£100 000"; education level coded as "College or University degree," "A levels/AS levels or equivalent," "O levels/GCSEs or equivalent," or "others;" and employment categorised as "working-in paid employment or self-employed", "unemployed/retired," or "other" (looking after the home and/or family, unable to work because of sickness or disability, doing unpaid or voluntary work, full or part-time student).

The Townsend deprivation index, validated for use in the UK-based population, was used as a measure of socio-economic status. [29] This measure combines census data on housing, employment, social class, and car availability based on the postal code of participants. The index was categorised into quintiles based on the baseline sample, with the least deprived (quintile 1) to the most deprived (quintile 5). This Townsend deprivation category was used as the measure of socio-economic deprivation (independent variable). Shift work, poor health, obesity and depressive symptoms were considered as the mediators, and poor sleep was considered as the outcome in mediation analysis.

Based on a priori evidence, three groups of variables were included in the analysis (i) lifestyle, (ii) shift work; and (iii) health factors. [30-33]

Lifestyle factors: Adapting a previously published method, we used data on lifestyle factors to generate a lifestyle score using smoking status, alcohol intake, television viewing time, computer hours (for recreational purposes), physical activity, tea and coffee consumption, fruit and vegetable intake, oily fish intake, and red and processed meat intake. This scale was initially used for assessing lifestyle in a large population study in Australia [34] and later adapted for use with the UK Biobank data. [35]

Dietary intake was evaluated through the 24-hour dietary recall to assess the consumption of processed meat, poultry, beef, lamb, pork, oily fish, non-oily fish, fresh fruit, dried fruit, raw vegetables, and cooked vegetables. Caffeine intake was assessed based on daily coffee intake, categorised as "none" "1-3 cups/day", and">3 cups/day." Tea intake was assessed based on daily tea intake, categorised as "none" "1-6 cups/day", and ">6 cups/day" [36]. Physical activity (PA) data were analysed following the International Physical Activity Questionnaire (IPAQ) scoring protocol with total physical activity computed as the sum of walking, moderate and vigorous activity in a previous week.[37] Based on standard scoring criteria, PA level was categorised as "inadequate (<150 min per week of moderate-intensity physical activity or <75 min per week of vigorous activity)" and "adequate (>150 min per week of moderate-intensity physical activity or >75 min per week of vigorous activity) [26]. Alcohol consumption was classified as "daily," "1-4 times/week," "sometimes," and "never." Smoking status was categorised as "current," "former," and "never smoked". Sedentary behaviour was computed as the sum of daily hours spent watching TV or working on the computer for recreational purposes. Each variable was dichotomised (0 points if not at risk, 1 point if at risk). Participants received 1 point for each unhealthy category (current smoker; alcohol consumed daily or almost daily; <150 min per week of moderate-intensity physical

activity or <75 min per week of vigorous activity; >4 h per day of sedentary behaviour; <400 g of fruits and vegetables per day; < 1 portion of oily fish per week; > three portions of red meat per week; > one portion of processed meat per week, >6 cups of tea per day and >4 cups of coffee per day). The criteria to categorise at-risk lifestyle are based on national dietary guidelines and available evidence and are reported in detail elsewhere. [34] Participants' scores were summed to create an unhealthy lifestyle index with a minimum score of zero, indicating the most healthy lifestyle, and a maximum score of nine, indicating the least healthy lifestyle. To examine the association of lifestyle with health outcomes, participants were classified into three categories according to their lifestyle score. Participants who scored 0-2 were classed as "healthy"; those who scored 3-5 were classed as "moderately healthy"; and those who scored 6+ were classed as "least healthy". [35]

Shift work: Participants were asked whether their main job involved night shifts, defined as "a work schedule that involves working through the normal sleeping hours." Response options were "rarely/never," "sometimes," and "often."

Health factor: BMI was categorised as "underweight <18.5 kg/m^2", "normal-between ≥18.5 and <25 kg/m^2, "overweight-between ≥25 and <30 kg/m^2," and "obese ≥30 kg/m^2". [38] Depressive symptoms were explored by asking " "Over the last two weeks, how often have you been bothered by feeling down, depressed, or hopeless" about the recent feeling of depression with responses "not at all," "several days," and "often." Satisfaction with overall health was assessed by asking, "in general how would you rate your overall health" the responses were categorised as "excellent/good," "fair," and "poor."

Statistical analysis

Latent trajectory analysis

We used group-based multi-trajectory modelling (GBMTM) to identify latent groups of individuals and profile the characteristics of individuals within the clusters. [39] GBMTM

combines features of latent class analysis and multilevel modelling to study variation in longitudinal outcomes and groups of individuals into meaningful clusters. [39] GBMTM allows simultaneous modelling of multiple measures of the same underlying construct and identifies group trajectories presented by a finite set of different polynomial functions of time, using maximum likelihood estimation. [39] The GBMTM model provides three pieces of information: the number of groups that best describe the data, the average trajectory for each group, and a probability estimate that each individual belongs to a particular group. [40] Each individual is assigned to the set of trajectories for which he/she has the highest probability to belong. The GBMTM handles missing values by fitting the model using maximum likelihood estimation, generating asymptotically unbiased parameters estimates, assuming that the data are missing at random. [41]

A two-stage approach was used to identify the best fitting model. We started the GBMTM with a single-group model to identify if a multi-trajectory approach provided a better fit to the data than a single group. Models with one to six groups were run to compare the Bayesian information criteria (BIC) of each model. A model with larger BIC indicated a better fitting model. [42] Next, polynomial terms were fitted in the models. The model providing the best fit and favouring parsimony was selected as the final model. The chosen model quality was verified according to these recommended criteria: the average posterior probabilities (AvePP) for each subgroup (≥0.7), the odds of correct classification (≥5), and adequate sample size in each group. [42] Sleep trajectories were modelled using the censored normal distribution using the information from only those participants who had data collected at two or more time points. We implemented GBMTM with the *Traj* plugin in Stata. [43]

Multinomial logistic regression (complete case analysis) was used to identify the independent impact of each variable on sleep trajectories. Firstly, collinearity among variables was examined by computing the variance inflation factor (VIF). Bi-variable associations between

sleep trajectories and baseline risk factors were examined to screen potential risk factors. Risk factors with p-value <0.20 were retained in the multivariable model. The p-value of 0.05 was adopted as a significance threshold for multinomial regression. All statistical analyses were undertaken using Stata IC 15.0 (Stata Statistical Software, College Station, Tx, USA). Based on a counterfactual framework, mediation analysis (logistic regression) was conducted to assess the role of mediators, i.e., shift work, poor health, obesity and depressive symptoms on the association between socio-economic deprivation and poor sleep. To capture socio-economic deprivation, we used the Townsend deprivation scores as it is a robust measure which concurrently captures multiple indicators of deprivation, e.g., unemployment, non-home ownership, household overcrowding etc. People in the most deprived conditions, i.e., positive values of the index (indicating high material deprivation) were captured the quintiles 4 & 5 of the index, and therefore a binary variable was created to capture socio-economic deprivation for mediation analysis. Using the Paramed package in STATA, we estimated the natural direct effects (NDEs), controlled direct effects (CDEs) and natural indirect effects (NIEs) of mediators after controlling for potential risk factors. [44] Furthermore, the Paramed program can also help in estimating NDEs and NIEs in the presence of risk factor-mediator interaction. [45] Each mediator was assessed separately, to estimate its unique direct and indirect effect after adjusting for other factors.

Results

In this study, sleep trajectories were examined by analysing data collected over three-time points for 41,094 participants (51.9% females) with a median age of 57 years (IQR 50-62 years). The majority of respondents were of white ethnicity (97.3%), had attended college or university (47.0%), were in a paid employment (65.2%), and reported gross annual household income as >£52 000 (31.6%). A considerable proportion of the participants were overweight/obese (62.8%), and some participants (14.6%) reported that their job involved

shift work. The mean follow-up time from the first assessment to the third assessment was 6.8 years.

Sleep Trajectories

As the number of groups increased, BIC increased as well ($BIC_{one\ group}$=- -338718.11, $BIC_{two\ groups}$=- 327124.29, $BIC_{three\ groups}$=-323747.26, $BIC_{four\ groups}$=-320563.10, $BIC_{five\ groups}$=-317517.05, $BIC_{six\ groups}$=- 315502.98, $BIC_{seven\ groups}$=-314127.00), however, with the increase in the number of groups, the group sizes became smaller, and groups lost distinctiveness. The three group model ($BIC_{theree\ groups}$=- 323637.55), with polynomial terms for sleep problems (111), sleep duration (112) and daytime sleepiness (122) offered three distinct classes with the smallest group of adequate size (27.6%) to allow further analysis. The three-group model also offered AvePP in the range of 0.82 to 0.92. The odds of correct classification based on the posterior probabilities of group membership were over 5.0 for all three groups indicating that the model had good assignment accuracy and provided an adequate fit to the data (Table-1). The model with the three classes also best met the theoretical criteria for identifying the optimal model. Therefore, we chose to select the 3-class model with trajectories categorised as 1) healthy sleepers; 2) borderline poor sleepers, and 3) poor sleepers. The three-group model we chose is shown in Figure-1.

The participants in the healthy sleeping trajectory, which was also the largest trajectory (n = 16,770; 40.8%) reported consistent sleep duration across all three time point (T1: 7.24 hrs; T2: 7.32 hrs T3:7.28 hrs) that was within the range of recommended sleep for age (7 to 9 hours per day) [46]. Compare with all other trajectories, participants in this group scored lower sleep problems (T1: 1.09; T2: 1.02; T3:1.10) and daytime sleepiness (T1: 0.36; T2: 0.01; T3:0.0) across all three-time points. Therefore, this group was labelled as “healthy sleepers.” The participants in the second group (n = 12,994; 31.6%) also reported consistent sleep duration across all three-time points that were within the recommended sleep duration range

(T1: 7.76 hrs; T2: 7.97 hrs T3:7.75 hrs). However, compared with the healthy sleeping trajectory, participants in the borderline poor sleeping group scored higher for sleep problems (T1: 1.26; T2: 1.29; T3:1.30) and daytime sleepiness (T1: 0.96; T2:1.31; T3:2.00). Therefore, this group was labelled as "borderline poor sleepers."

The participants in the third group (n = 11,330; 27.6%) reported consistent short sleep duration across all three-time points (T_1: 6.38 hrs; T_2: 6.29 hrs T_3:6.22 hrs). The participants in this trajectory started and continued with high scores for sleep problems (T_1: 1.62; T_2: 1.57; T_3:1.60) and daytime sleepiness (T_1: 1.21; T_2: 1.42; T_3:1.98) than the borderline healthy sleeping group. Therefore, this group was labelled as "poor sleepers."

Association between baseline socio-demographic, lifestyle, occupation and health conditions and sleep trajectories

Results from bi-variable association analysis indicated that baseline socio-economic status, unhealthy lifestyle and poor health were significantly linked with poor sleeping trajectories (Supplementary Table-1). Collinearity analysis led to the removal of the employment variable from the groups of risk factors. For the remaining risk factors, the variance inflation factor (VIF) ranged from 1.05 to 1.19 indicated that collinearity was absent. Therefore, all the remaining risk factors were used in regression analysis. Factors identified through bivariate association were entered in separate multinomial regression models to assess the impact of each group of variables, i.e., socio-demographic variables (age, gender, ethnicity, education, income, Townsend score); lifestyle; shift work and health conditions (depressive symptoms, obesity and poor health) on poor sleep (Table-2). The multinomial regression results are presented for the factors associated with being poor sleeper group relative to healthy sleepers, the risk factors being borderline poor sleepers relative to healthy sleepers group are presented in supplementary table-2.

Multinomial regression results suggest that older adults, females and ethnic minorities had a higher risk of being in poor sleeper group, relative to healthy sleepers, given all other variables in the model are held constant (Model-1). Socio-economic disadvantages, i.e., low household income, ethnic minority status (50% increased risk) and being in the most deprived category on the Townsend deprivation measure (nearly 40% increased risk) significantly increased the risk of being in the poor sleep group (Model-1). Unhealthy lifestyle (82% increased risk) (Model-2), and a shift based job (58% increased risk) (Model-3) also increased the risk of being in the poor sleeping group. Health conditions (Model-4), such as obesity (53% increased risk) were associated with increased risk of being in the poor sleeping group, while participants with excellent health had significantly reduced risk (87% reduced risk) of being in the poor sleeping category. Notably, often experiencing depression resulted in a more than 5-fold risk of being in the poor sleep group relative to having no depression, and even depression on several days a week was associated with a more than 3-fold risk. Potential risk factors, e.g., lifestyle (Model-5), shift work (Model-6), and health conditions (Model-7) were separately added to regression models to identify the factors affecting the association between socio-economic disadvantages and poor sleep. Despite adding lifestyle (Model-5) and then shift work (Model-6) to the model, the effect sizes for the impact of socio-economic disadvantage (ethnicity, income and Townsend score) on poor sleep did not change significantly. These findings suggest that neither lifestyle nor shift work is affecting the associations between socio-economic disadvantage and poor sleep.

The addition of variables related to health conditions attenuated the impact of socio-demographic factors on the risk of being in the poor sleep group. Importantly, lower education, which was a significant factor associated with poor sleep in previous models, became non-significant upon addition of variables related to health conditions (Model-7). In the final model (Model-8), all variables were entered concurrently to identify the main factors

associated with poor sleep. While concurrent analysis of variables attenuated the impact of socio-demographic disadvantages, unhealthy lifestyle (70% increased risk), shift work (40%), depressive symptoms (still more than 5-fold for those often depressed), obesity (nearly 50% increased risk) and poor health were still significantly linked with increased risk of being in the poor sleep group.

Results from the final multinomial regression (complete case analysis, n= 15,330) suggested that the association between socio-economic disadvantage and poor sleep is potentially mediated through health conditions. Mediation analyses (Table-3) indicate that depressive symptoms are the primary mediator of poor sleep in socio-economically deprived communities, with 32.8% of the total effect of socio-economic deprivation on poor sleep found to be mediated through depressive symptoms alone. Poor health and obesity were other significant mediators of the association between socio-economic deprivation and poor sleep as 7.5% of the total effect were found to be mediated through both poor sleep and obesity.

Discussion

This is the first study empirically to identify trajectories of sleep for a very large adult population longitudinally. Of particular note is that over six years of follow-up, these trajectories are stable. This indicates that poor sleep is a relatively stable phenomenon; therefore, early intervention is required for sleep improvement, rather than assuming that poor sleep is a transient problem that will remit spontaneously.

It is important to highlight that participants in the borderline poor sleeping group (around 1/3rd of the cohort), who reported sleeping for the recommended duration (around 30 minutes longer than healthy sleepers), still experienced a considerable burden of sleep problems and daytime sleepiness. This indicates that if we wish to optimise health outcomes relating to sleep, we need to consider more than just healthy sleep duration, focusing on the whole spectrum of sleep health issues to improve early diagnosis and management of poor sleep.

A second key finding from this analysis is that the relationships between education, minority status and sleep, attenuates when the experience of depressive symptoms is added to the multivariable models. This seems to be suggesting that the experience of depressive symptoms mediates the relationship between indices of socio-economic status (income, education) or social determinants of health and poor sleep. This could be interpreted to indicate that socio-economic disadvantages increase the risk of depressive symptoms, which in turn further increases the risk of poor sleep in people from socio-economically disadvantaged groups. However, we also need to acknowledge that people may be experiencing depressive symptoms as a consequence of persistent poor quality sleep. It is easy to see how disadvantage, resulting in the daily experience of discrimination for minority groups, and financial, food and housing security may impact on individuals' sleep quality. This could be either through material mechanisms (occupational density, thermal comfort, and unsafe environments) or through the psychological impact of insecurity and discrimination. Functional neuroimaging findings have further supported the strong link and a potential dynamic bidirectional relationship between major sleep disorders and affective (internalising) disorders.[47] Also, the findings of a recent meta-analysis have suggested that even acute sleep deprivation may lead to profound early changes in the critical brain nodes where multiple cortex computations enable our choosing of an action to occur.[47] The authors argued that these initial brain circuitry changes might initiate a cascade leading to longer-term maladaptive internalising coping skills, dysmetria of thought and action, which are recognised risk factors for depression and anxiety disorders and further sleep issues and disorders. [47] Regardless of the mechanism, this highlights the close interaction between sleep and emotional well-being [48] and the need to address both issues, rather than focusing on one in the hope it will address both. In this context, it is unlikely we will kill two birds with one stone.

The third point of note in this large longitudinal study is that men are less likely to report poor sleep than women in all models, even after adjusting for disadvantage. While one might argue that this is an artefact of measurement non-invariance between men and women, objective sleep studies confirm these gender differences, supporting their importance.[49] Considering the age of the participants included in this study, the high prevalence of poor sleep in this cohort could be explained in the context of the menopausal transition. Though the lack of information on menopausal status is a limitation of our work, nonetheless, previous studies have established the link between hot flashes, changes in follicle-stimulating hormone and poor sleep. [50] Therefore high prevalence of poor sleep in females can be attributed to menopausal transition and menopausal status. However, we also need to consider whether this represents a reporting bias. This is further supported by the fact that, in the analysis of actigraphy data from a subset of the UK Biobank, women slept longer than men. [51] This indicates that we need to investigate relationships between objective and subjective sleep data differentially for men and women and, where intervention resources are limited, perhaps women, specifically those from ethnic minority groups should be the highest priority.

While we acknowledge that the data reported here are based on self-reported sleep data, we have used more than just self-reported sleep duration to create the sleep trajectories. It should be noted that the UK Biobank cohort is not representative of the general UK population, which limits the generalisability of our findings, furthermore being an observational study; the results are prone to bias. However, we used a sound analytical approach and adjusted for a range of established risk factors in our analyses. We used multiple indicators of sleep and daytime functioning issues to identify sleep trajectories. Given the size of the sample, and its longitudinal design, this use of multiple indicators of sleep problems gives a degree of

robustness to the study. A key next step is to determine if these sleep trajectories align with more objective indicators of individuals' sleep patterns. A sub-group of the UK Biobank agreed to wear actigraphs for seven days. The data from this sample of 103,000 individuals have now been processed, and derived variables for sleep have just been made available.[52] While a direct comparison of this cross-sectional data has yet to be undertaken, Jones and colleagues report matching directional associations between self-reported data and actigraph data in relation to genetic variants associated with sleep parameters. Thus, our next step will be to see to what extent self-reported data and actigraphy data align for this cohort.

In conclusion, we have identified three stable sleep trajectories for a large cohort of UK individuals and identified risk factors and mediators of poor, which could be used for identifying individuals for early intervention.

Acknowledgements

This research was conducted using the UK Biobank resource. The authors would like to thank the UK Biobank participants and investigators for making this study possible.

Financial disclosures

None.

Non-financial disclosures

None

References

1. Liu Y, Wheaton AG, Chapman DP, Cunningham TJ, Lu H, Croft JB. Prevalence of Healthy Sleep Duration among Adults--United States, 2014. MMWR Morbidity and mortality weekly report. 2016; 65 (6): 137-141.
2. Medic G, Wille M, Hemels ME. Short- and long-term health consequences of sleep disruption. Nature and science of sleep. 2017; 9: 151-161.
3. Lund HG, Reider BD, Whiting AB, Prichard JR. Sleep patterns and predictors of disturbed sleep in a large population of college students. The Journal of adolescent health : official publication of the Society for Adolescent Medicine. 2010; 46 (2): 124-132.
4. Kang J-M, Lee JA, Jang J-W, Kim YS, Sunwoo S. Factors associated with poor sleep quality in primary care. Korean J Fam Med. 2013; 34 (2): 107-114.

5. Morrison DN, McGee ROB, Stanton WR. Sleep Problems in Adolescence. Journal of the American Academy of Child & Adolescent Psychiatry. 1992; 31 (1): 94-99.
6. Waters KA, Suresh S, Nixon GM. Sleep disorders in children. The Medical journal of Australia. 2013; 199 (8): S31-35.
7. Jarrin DC, McGrath JJ, Drake CL. Beyond sleep duration: distinct sleep dimensions are associated with obesity in children and adolescents. Int J Obes (Lond). 2013; 37 (4): 552-558.
8. Fatima Y, Doi SA, Mamun AA. Sleep quality and obesity in young subjects: a meta-analysis. Obesity reviews : an official journal of the International Association for the Study of Obesity. 2016; 17 (11): 1154-1166.
9. Salo P, Vahtera J, Ferrie JE, et al. Trajectories of sleep complaints from early midlife to old age: longitudinal modeling study. Sleep. 2012; 35 (11): 1559-1568.
10. Lanza ST, Rhoades BL, Greenberg MT, Cox M. Modeling multiple risks during infancy to predict quality of the caregiving environment: Contributions of a person-centered approach. Infant Behavior and Development. 2011; 34 (3): 390-406.
11. Chen J, Patel SR, Redline S, et al. Weekly sleep trajectories and their associations with obesity and hypertension in the Hispanic/Latino population. Sleep. 2018; 41 (10).
12. Plancoulaine S, Reynaud E, Forhan A, Lioret S, Heude B, Charles MA. Night sleep duration trajectories and associated factors among preschool children from the EDEN cohort. Sleep medicine. 2018; 48: 194-201.
13. Laursen BH, E. Person-Centered and Variable-Centered Approaches to Longitudinal Data. Merrill-Palmer Quarterly. 2006; 50 (2): 377-389.
14. Howard MC, Hoffman ME. Variable-Centered, Person-Centered, and Person-Specific Approaches: Where Theory Meets the Method. Organizational Research Methods. 2017; 21 (4): 846-876.
15. Porcu M, Giambona F. Introduction to Latent Class Analysis With Applications. The Journal of Early Adolescence. 2016; 37 (1): 129-158.
16. Halfon N, Hochstein M. Life course health development: an integrated framework for developing health, policy, and research. Milbank Q. 2002; 80 (3): 433-iii.
17. Park J, Yu H-T. Recommendations on the Sample Sizes for Multilevel Latent Class Models. Educational and Psychological Measurement. 2017; 78 (5): 737-761.
18. Grandner MA, Jackson NJ, Izci-Balserak B, et al. Social and Behavioral Determinants of Perceived Insufficient Sleep. Frontiers in Neurology. 2015; 6 (112).
19. Johnson DA, Jackson CL, Williams NJ, Alcantara C. Are sleep patterns influenced by race/ethnicity - a marker of relative advantage or disadvantage? Evidence to date. Nature and science of sleep. 2019; 11: 79-95.
20. Vishnu A, Shankar A, Kalidindi S. Examination of the Association between Insufficient Sleep and Cardiovascular Disease and Diabetes by Race/Ethnicity. International journal of endocrinology. 2011; 2011: 789358.
21. Williamson AA, Mindell JA. Cumulative socio-demographic risk factors and sleep outcomes in early childhood. Sleep. 2019; 43 (3).
22. Bartel KA, Gradisar M, Williamson P. Protective and risk factors for adolescent sleep: a meta-analytic review. Sleep Med Rev. 2015; 21: 72-85.
23. Jackson CL, Walker JR, Brown MK, Das R, Jones NL. A workshop report on the causes and consequences of sleep health disparities. Sleep. 2020; 43 (8).
24. Sekine M, Chandola T, Martikainen P, Marmot M, Kagamimori S. Work and Family Characteristics as Determinants of Socio-economic and Sex Inequalities in Sleep: The Japanese Civil Servants Study. Sleep. 2006; 29 (2): 206-216.

25. Lallukka T, Sares-Jäske L, Kronholm E, et al. Socio-demographic and socio-economic differences in sleep duration and insomnia-related symptoms in Finnish adults. BMC Public Health. 2012; 12 (1): 565.
26. Allen NE, Sudlow C, Peakman T, Collins R. UK biobank data: come and get it. Science translational medicine. 2014; 6 (224): 224ed224.
27. UK Biobank. UK Biobank: Protocol for a large-scale prospective epidemiological resource http://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf.
28. Ginneken JV, Groenewold G. A single- vs. multi-item self-rated health status measure: a 21-country study. The Open Public Health Journal. 2012; 5: 1-9.
29. Townsend P. Poverty in the United Kingdom In. London: Allen Lane and Penguin Books'; 1979.
30. Grandner MA, Patel NP, Gehrman PR, et al. Who gets the best sleep? Ethnic and socio-economic factors related to sleep complaints. Sleep medicine. 2010; 11 (5): 470-478.
31. Mezick EJ, Matthews KA, Hall M, et al. Influence of race and socio-economic status on sleep: Pittsburgh SleepSCORE project. Psychosom Med. 2008; 70 (4): 410-416.
32. Ohida T, Kamal AM, Uchiyama M, et al. The influence of lifestyle and health status factors on sleep loss among the Japanese general population. Sleep. 2001; 24 (3): 333-338.
33. Jehan S, Zizi F, Pandi-Perumal SR, et al. Shift Work and Sleep: Medical Implications and Management. Sleep Med Disord. 2017; 1 (2): 00008.
34. Ding D, Rogers K, van der Ploeg H, Stamatakis E, Bauman AE. Traditional and Emerging Lifestyle Risk Behaviors and All-Cause Mortality in Middle-Aged and Older Adults: Evidence from a Large Population-Based Australian Cohort. PLOS Medicine. 2015; 12 (12): e1001917.
35. Foster HME, Celis-Morales CA, Nicholl BI, et al. The effect of socio-economic deprivation on the association between an extended measurement of unhealthy lifestyle factors and health outcomes: a prospective analysis of the UK Biobank cohort. The Lancet Public Health. 2018; 3 (12): e576-e585.
36. Bradbury KE, Murphy N, Key TJ. Diet and colorectal cancer in UK Biobank: a prospective study. International Journal of Epidemiology. 2019.
37. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. Medicine and science in sports and exercise. 2003; 35 (8): 1381-1395.
38. World Health Organization. *Obesity: preventing and managing the global epidemic.* Switzerland 3–5 June 1997.
39. Nagin DS, Jones BL, Passos VL, Tremblay RE. Group-based multi-trajectory modeling. Statistical methods in medical research. 2018; 27 (7): 2015-2023.
40. Nagin DS. Group-Based Trajectory Modeling: An Overview. Annals of Nutrition and Metabolism. 2014; 65 (2-3): 205-210.
41. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annual review of clinical psychology. 2010; 6: 109-138.
42. D. N. *Group-based modeling of development.* Cambridge: Harvard University Press; 2005.
43. Jones BL, Nagin DS. A Note on a Stata Plugin for Estimating Group-based Trajectory Models. Sociological Methods & Research. 2013; 42 (4): 608-613.
44. *"PARAMED: Stata module to perform causal mediation analysis using parametric regression models* [computer program]. Boston College Department of Economics; 2013.

45. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychological methods. 2013; 18 (2): 137-150.
46. Hirshkowitz M, Whiton K, Albert SM, et al. National Sleep Foundation’s sleep time duration recommendations: methodology and results summary. Sleep Health: Journal of the National Sleep Foundation. 1 (1): 40-43.
47. Khazaie H, Veronese M, Noori K, et al. Functional reorganisation in obstructive sleep apnoea and insomnia: A systematic review of the resting-state fMRI. Neurosci Biobehav Rev. 2017; 77: 219-231.
48. Vandekerckhove M, Cluydts R. The emotional brain and sleep: An intimate relationship. Sleep medicine reviews. 2010; 14 (4): 219-226.
49. Mallampalli MP, Carter CL. Exploring sex and gender differences in sleep health: a Society for Women's Health Research Report. J Womens Health (Larchmt). 2014; 23 (7): 553-562.
50. Baker FC, de Zambotti M, Colrain IM, Bei B. Sleep problems during the menopausal transition: prevalence, impact, and management challenges. Nature and science of sleep. 2018; 10: 73-95.
51. Zhu G, Catt M, Cassidy S, et al. Objective sleep assessment in >80,000 UK mid-life adults: Associations with socio-demographic characteristics, physical activity and caffeine. PLOS ONE. 2019; 14 (12): e0226220.
52. Jones SE, van Hees VT, Mazzotti DR, et al. Genetic studies of accelerometer-based sleep measures yield new insights into human sleep behaviour. Nature Communications. 2019; 10 (1): 1585.

Figure-1 Caption

Figure-1: Sleep health trajectories in the UK Biobank cohort based on sleep duration, sleep problems and daytime sleepiness data obtained from a sample of 41,094 participants (2006-2016)

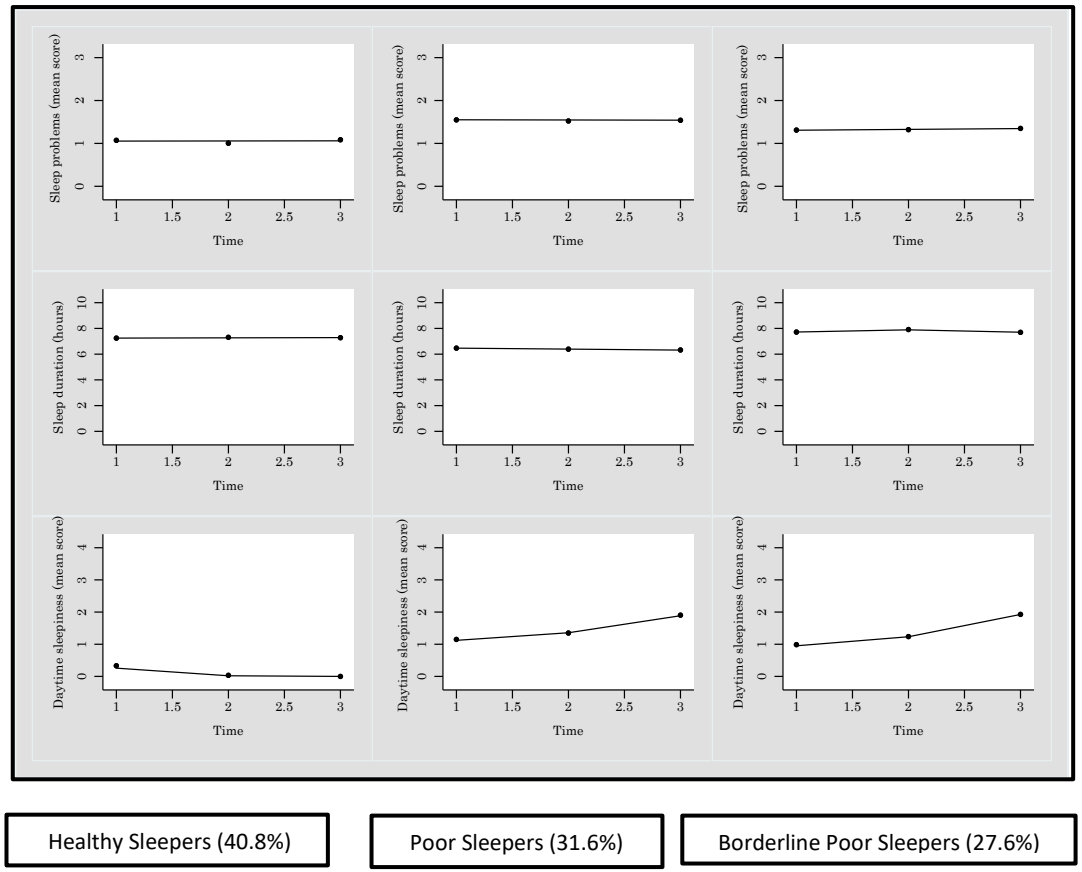

Figure-1: Sleep health trajectories in the UK Biobank cohort based on sleep duration, sleep problems and daytime sleepiness data obtained from a sample of 41,094 participants (2006-2016)

Table-1: Group Membership Probabilities, and Odds of Correct Classification for three group-based latent trajectories of sleep in 41,094 participants of the UK Biobank cohort (2006-2016).

Sleep Trajectory	Trajectory Size n(%)	Mean Probability of Group Membership	Odds of Correct Classification	Difference between proportion assigned and the estimated probability (P-π)
Healthy Sleepers	16,770 (40.8)	0.92	337.5	0.0009
Borderline Poor Sleepers	12,994 (31.6)	0.83	177.1	0.001
Poor Sleepers	11,330 (27.6)	0.82	189.7	0.0002

Table-2: Predictors of poor sleep in the UK Biobank cohort (2006-2016), using healthy sleep trajectory as the reference group

Baseline Predictors	Model-1		Model-2		Model-3		Model-4		Model-5		Model-6		Model-7		Model-8	
	RRR	95%CI	RRR	95%CI	RRR	95%CI	RRR	95%CI	RRR	95%CI	RRR	95%CI	RRR	95%CI	RRR	95%CI
Age	**0.98**	**0.97-0.99**							**0.98**	**0.97-0.99**	**0.98**	**0.97-0.99**	**0.98**	**0.98-0.99**	0.99	0.98-1.00
Gender																
Female (Ref)																
Male	**0.89**	**0.85-0.94**							**0.84**	**0.80-0.89**	**0.89**	**0.83-0.94**	**0.88**	**0.92-.94**	**0.81**	**0.75-0.88**
Ethnicity																
White (Ref)																
Ethnic minorities	**1.55**	**1.32-1.82**							**1.62**	**1.38-1.90**	**1.53**	**1.28-1.83**	**1.31**	**1.04-1.64**	**1.39**	**1.08-1.78**
Townsend Score																
Quartile-1 (Ref-Least deprived)																
Quartile-2	1.06	0.98-1.14							1.05	0.97-1.14	1.08	0.99-1.19	0.99	0.90-1.10	1.02	0.90-1.14
Quartile-3	1.07	0.99-1.16							1.07	0.99-1.16	1.07	0.97-1.17	1.04	0.94-1.15	1.03	0.91-1.16
Quartile-4	**1.23**	**1.14-1.34**							**1.22**	**1.12-1.32**	**1.25**	**1.13-1.38**	1.10	0.99-1.23	**1.14**	**1.01-1.29**
Quartile-5 (Most deprived)	**1.39**	**1.26-1.52**							**1.37**	**1.25-1.51**	**1.40**	**1.26-1.56**	**1.2**	**1.10-1.40**	**1.28**	**1.11-1.48**
Annual household income (£)																
<18,000 (Ref)																
18,000 to 30,999	**0.83**	**0.76-0.92**							**0.84**	**0.76-0.92**	0.87	0.75-1.01	0.94	0.83-1.07	0.94	0.77-1.14
31,000 to 51,999	**0.76**	**0.69-0.83**							**0.76**	**0.70-0.84**	**0.75**	**0.65-0.86**	0.90	0.80-1.02	0.85	0.71-1.03
>52,000	**0.63**	**0.57-0.69**							**0.63**	**0.58-0.70**	**0.64**	**0.56-0.74**	**0.84**	**0.74-0.96**	**0.83**	**0.69-0.99**
Education																
College/Uni (Ref)																
A levels/AS level or equivalent	1.05	0.97-1.14							1.04	0.96-1.12	1.06	0.96-1.16	0.98	0.88-1.09	1.03	0.91-1.16
O levels/GCSEs or equivalent	**1.14**	**1.07-1.22**							**1.13**	**1.06-1.20**	**1.13**	**1.04-1.23**	1.05	0.96-1.15	1.06	0.95-1.17
Others	**1.12**	**1.03-1.22**							**1.11**	**1.02-1.21**	**1.11**	**1.01-1.24**	1.05	0.94-1.18	1.03	0.89-1.18
Lifestyle**																
Healthy (Ref)																
Moderate Healthy			**1.29**	**1.23-1.36**					**1.33**	**1.26-1.40**					**1.26**	**1.15-1.37**
Unhealthy			**1.82**	**1.60-2.09**					**1.94**	**1.67-2.24**					**1.73**	**1.37-2.17**
Shift work																
Never (Ref)																
Sometimes					**1.27**	**1.14-1.43**					**1.15**	**1.02-1.30**			1.09	0.92-1.28
Usually/always					**1.58**	**1.42-1.75**					**1.42**	**1.26-1.60**			**1.42**	**1.21-1.66**
Depressive symptoms																
Not at all (Ref)																
Several days							**3.66**	**3.35-3.98**					**3.56**	**3.24-3.91**	**3.51**	**3.15-3.91**
Often							**5.70**	**4.59-7.08**					**5.43**	**4.27-6.91**	**5.39**	**4.05-7.17**
Overweight/Obesity																
Normal (Ref)																
Overweight							**1.11**	**1.04-1.19**					**1.18**	**1.09.1.28**	**1.14**	**1.04-1.25**

Obese	**1.47**	**1.34-1.60**	**1.54**	**1.40-1.70**	**1.48**	**1.31-1.66**
Overall health						
Poor (Ref)						
Excellent/good	**0.13**	**0.09-0.18**	**0.14**	**0.10-0.21**	**0.20**	**0.12-0.32**
Fair	**0.36**	**0.26-0.50**	**0.40**	**0.28-0.59**	**0.50**	**0.31-0.83**

RR: Relative Risk Ratio, Significant predictors are highlighted in bold

* lifestyle variable includes smoking, alcohol consumption, physical activity, TV hours, computer hours, tea, coffee, fruit, vegetables, fish, processed meat, red meat consumption

Model-1 concurrent consideration of all sociodemographic variables

Model-2 lifestyle variable

Model-3 shiftwork

Model-4 concurrent consideration of all health related variables (depressive symptoms, obesity and health status)

Model-5 all sociodemographic variables+ lifestyle

Model-6 all sociodemographic variables+ shiftwork

Model-7 all sociodemographic variables+ all health related variables

Model-8 all sociodemographic variables+ lifestyle + shiftwork +all health related variables

Table-3: The contribution of health conditions in mediating the association between socio-economic disadvantages and poor sleep, based on UK Biobank cohort (2006-2016) data

Mediator	Estimate	SE	95%CI
Mediation by poor health*			
Controlled Direct Effect	1.14	0.04	1.07-1.22
Natural Direct Effect (NDE)	1.14	0.04	1.07-1.22
Natural Indirect Effect (NIE)	1.01	0.008	1.01-1.03
Total Mediation[#]	7.5%		
Mediation by obesity**			
Controlled Direct Effect	1.14	0.04	1.07-1.22
Natural Direct Effect (NDE)	1.14	0.04	1.07-1.22
Natural Indirect Effect (NIE)	1.01	0.003	1.00-1.01
Total Mediation[#]	7.5%		
Mediation by depression***			
Controlled Direct Effect	1.14	0.04	1.07-1.22
Natural Direct Effect (NDE)	1.14	0.04	1.07-1.22
Natural Indirect Effect (NIE)	1.06	0.01	1.03-1.07
Total Mediation[#]	32.8%		

Mediation models are based on socio-economic deprivation (Townsend deprivation quartile 4-5) as the exposure, poor sleeping trajectory as the outcome and above-mentioned variables as the mediators

* adjusted for age, gender, lifestyle, shift work, obesity, depression

** adjusted for age, gender, lifestyle, shift work, health, depression

*** adjusted for age, gender, lifestyle, shift work, obesity, health

[#]Calculation of proportion mediated: (NDE*(NIE-1)) / (NDE*NIE -1)

Sleep trajectories and mediators of poor sleep: Findings from the longitudinal analysis of 41,094 participants of the UK Biobank cohort

Yaqoot Fatima[1, 2*], Romola S Bucks[3], Abdullah A Mamun[1], Isabelle Skinner[2], Ivana Rosenzweig [4], Guy Leschziner[5, 6], Timothy C Skinner[2], [7, 8]

1. Institute for Social Science Research, University of Queensland, Brisbane, Australia
2. Centre for Rural and Remote Health, James Cook University, Mount Isa, Australia
3. School of Psychological Science, University of Western Australia, Perth, Australia
4. Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, United Kingdom
5. Sleep Disorders Centre, Guy's and St Thomas' NHS Foundation Trust, Guy's Hospital, London, United Kingdom
6. Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
7. University Department of Rural Health, La Trobe University, Bendigo, Australia
8. Institut for Psykologi, Center for Sundhed of Samfund, Københavns Universitet, Øster Farimagsgade, København K. Denmark

*Address for Correspondence

Dr Yaqoot Fatima

Institute for Social Science Research

University of Queensland, Indooroopilly, Queensland 4068, Brisbane, Australia

Email: fatima.yaqoot@uq.edu.au

Telephone: +61 7 344 31021

Supplementary Table 1: Characteristics of the overall sample (n=41,094) and sleep trajectory subgroups in the UK Biobank cohort (2006-2016).

Variable	Total Sample	Healthy Sleepers	Borderline Poor Sleepers	Poor Sleepers	p-value*
Age (years)	55.7 (7.59)	55.9 (7.34)	56.1 (7.84)	55.0 (7.60)	<0.001
Gender					
Female	21,339 (51.9)	8,328 (49.7)	6,924 (53.3)	6,087 (96.1)	<0.001
Male	19,755 (48.1)	8,442(50.3)	6,070 (46.7)	5,243 (3.9)	
Ethnicity					
White	39,852 (97.3)	16,351 (97.7)	12,651 (97.6)	10,850(27.2)	<0.001
Ethnic minorities**	1,113(2.72)	378 (2.3)	299 (2.4)	436 (39.2)	
Townsend Score					
Quartile-1 (Least deprived)	10,018 (24.4)	4,363 (26.0)	920 (24.3)	2,500 (22.1)	<0.001
Quartile-2	9,492 (23.1)	4,013 (23.4)	3,006 (23.2)	2,473 (21.8)	
Quartile-3	8,703 (21.2)	3,618 (21.6)	2,745 (21.2)	2,340 (20.7)	
Quartile-4	7,380 (18.0)	2,863 (17.1)	2,195 (17.9)	2,322 (19.4)	
Quartile-5 (Most deprived)	5,462 (13.3)	1,902 (11.4)	1,748 (13.5)	1,812(16.1)	
Annual household income (£)					
<18,000	5,274 (14.4)	1,798 (12.0)	1,890 (16.3)	1,586 (15.8)	<0.001
18,000-30,999	8,898 (24.2)	3,408 (22.7)	3,061 (26.4)	2,429 (24.2)	
31,000-51,999	10,919 (29.8)	4,468 (29.7)	3,438 (29.6)	3,013 (30.1)	
>52,000	11,596 (31.6)	5,369 (35.7)	3,223 (27.8)	3,004 (29.9)	
Education					
College/University	17,403 (47.0)	7,426(48.7)	5,483 (47.0)	4,495 (44.5)	<0.001
A Levels/AS Level or equivalent	5,105 (13.8)	2,105 (13.8)	1,613 (13.8)	1,387(13.7)	
O Levels /GCSEs or equivalent	9,888 (26.7)	3,847 (25.2)	3,121 (26.8)	2,920 (28.9)	

Others	4,610 (12.5)	1,877 (12.3)	1,439 (12.3)	1,294 (12.8)	
Lifestyle* **					
Most healthy	18,043 (43.9)	7,894 (47.1)	1,259 (43.0)	4,564 (40.3)	
Moderately healthy	21,677 (52.8)	8,426 (50.2)	6,960 (53.6)	6,291 (55.5)	<0.001
Unhealthy	1,374 (3.3)	450 (2.7)	449 (3.4)	475 (4.2)	
Shift Work					
Never	22,687 (85.4)	9,704 (87.2)	6,587 (85.3)	6,396 (82.8)	
Sometimes	1,776 (6.68)	686 (6.2)	514 (6.7)	576 (7.4)	<0.001
Usually/Always	2,105 (7.92)	728 (6.6)	621 (8)	756 (9.8)	
Depressive Symptoms					
Not At All	20,944 (79.3)	10,070 (90.1)	6,157 (74.6)	4,717 (67.5)	<0.001
Several Days	4,673 (17.7)	996 (8.8)	1,803 (21.8)	1,874 (26.8)	
Often	806 (3.05)	110 (0.1)	299 (3.6)	397 (5.7)	
Overweight/Obesity					
Normal	15,151 (37.1)	6,703 (40.2)	4,734 (36.7)	3,714 (33.0)	
Overweight	17,548 (43.0)	7,359 (44.2)	5,407 (41.9)	4782 (42.5)	<0.001
Obese	8,129 (19.9)	2,607 (15.6)	2,768 (21.4)	2,754 (24.5)	
Overall Health					
Excellent/Good	33,630 (82.1)	15,248 (91.1)	10,282 (79.4)	8,100 (71.8)	<0.001
Fair	6,407 (15.6)	1,421 (8.5)	2,280 (17.6)	2,706 (24.0)	
Poor	945 (2.31)	74 (0.4)	392 (3.0)	479 (4.2)	

Data are presented as %(n) or mean±sd., *Chi^2 test, **Black, Asian and mixed ethnicity, ***lifestyle variable includes smoking, alcohol consumption, physical activity, TV hours, computer hours, tea, coffee, fruit, vegetables, fish, processed meat, red meat consumption

Supplementary Table-2: Predictors of borderline poor sleep in the UK Biobank cohort (2006-2016), using healthy sleep trajectory as the reference group

Baseline Predictors	RRR	95%CI
Age	1.00	0.99-1.01
Gender		
Female (Ref)		
Male	**0.81**	**0.74-0.87**
Ethnicity		
White (Ref)		
Ethnic minorities	0.90	0.69-1.16
Townsend Score		
Quartile-1 (Ref-Least deprived)		
Quartile-2	1.00	0.90-1.12
Quartile-3	1.00	0.89-1.12
Quartile-4	1.00	0.89-1.13
Quartile-5 (Most deprived)	1.07	0.93-1.23
Annual household income (£)		
<18,000 (Ref)		
18,000 to 30,999	0.86	0.72-1.03
31,000 to 51,999	**0.78**	**0.66-0.93**
>52,000	**0.64**	**0.54-0.77**
Education		
College/Uni (Ref)		
A levels/AS level or equivalent	0.96	0.85-1.07
O levels/GCSEs or equivalent	**0.88**	**0.80-0.98**
Others	**0.88**	**0.77-1.00**
Lifestyle**		
Healthy (Ref)		
Moderate Healthy	**1.10**	**1.01-1.19**
Unhealthy	1.24	0.97-1.58

Shift work		
Never (Ref)		
Sometimes	1.02	0.86-1.19
Usually/always	**1.18**	**1.00-1.39**
Depressive symptoms		
Not at all (Ref)		
Several days	**2.66**	**2.39-2.97**
Often	**3.80**	**2.84-5.09**
Overweight/Obesity		
Normal (Ref)		
Overweight	0.99	0.89- 1.05
Obese	1.18	1.06-1.33
Overall health		
Poor (Ref)		
Excellent/good	**0.23**	**0.14-0.38**
Fair	**0.43**	**0.26-0.71**

RR: Relative Risk Ratio, Significant predictors are highlighted in bold