Journal of Affective Disorders 335 (2023) 83–94

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

Subjective and objective sleep and circadian parameters as predictors of depression-related outcomes: A machine learning approach in UK Biobank

Laura M. Lyall [a,b,*], Natasha Sangha [a,b], Xingxing Zhu [a], Donald M. Lyall [a], Joey Ward [a], Rona J. Strawbridge [a,c,d], Breda Cullen [a], Daniel J. Smith [a,b]

[a] School of Health and Wellbeing, University of Glasgow, Glasgow, UK
[b] Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK
[c] Health Data Research, UK
[d] Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institute, Stockholm, Sweden

ARTICLE INFO

Keywords:
Depression
Sleep
Circadian rhythms
Rest-activity
Inactivity
Postnatal depression

ABSTRACT

Background: Sleep and circadian disruption are associated with depression onset and severity, but it is unclear which features (e.g., sleep duration, chronotype) are important and whether they can identify individuals showing poorer outcomes.

Methods: Within a subset of the UK Biobank with actigraphy and mental health data (n = 64,353), penalised regression identified the most useful of 51 sleep/rest-activity predictors of depression-related outcomes; including case-control (Major Depression (MD) vs. controls; postnatal depression vs. controls) and within-case comparisons (severe vs. moderate MD; early vs. later onset, atypical vs. typical symptoms; comorbid anxiety; suicidality). Best models (of lasso, ridge, and elastic net) were selected based on Area Under the Curve (AUC).

Results: For MD vs. controls ($n_{(MD)}$ = 24,229; $n_{(control)}$ = 40,124), lasso AUC was 0.68, 95 % confidence interval (CI) 0.67–0.69. Discrimination was reasonable for atypical vs. typical symptoms ($n_{(atypical)}$ = 958; $n_{(typical)}$ = 18,722; ridge: AUC 0.74, 95 % CI 0.71–0.77) but poor for remaining models (AUCs 0.59–0.67). Key predictors across most models included: difficulty getting up, insomnia symptoms, snoring, actigraphy-measured daytime inactivity and lower morning activity (~8 am). In a distinct subset (n = 310,718), the number of these factors shown was associated with all depression outcomes.

Limitations: Analyses were cross-sectional and in middle-/older aged adults: comparison with longitudinal investigations and younger cohorts is necessary.

Discussion: Sleep and circadian measures alone provided poor to moderate discrimination of depression outcomes, but several characteristics were identified that may be clinically useful. Future work should assess these features alongside broader sociodemographic, lifestyle and genetic features.

1. Introduction

Depression has a severe impact on quality of life and daily functioning, particularly in the case of severe episodes, or where depression co-occurs alongside anxiety disorders or suicidal thoughts (Johnston et al., 2019). It is important to identify modifiable risk factors to target interventions aimed at preventing onset of depression, and among patients, to prevent more severe episodes and symptoms.

Sleep and circadian rhythm disturbances are important risk factors for depressive disorders. Sleep disturbances have commonly been viewed as a core symptom of major depression (MD), with up to 90 % of patients reporting sleep problems (Riemann et al., 2001). Prospective longitudinal studies have supported bidirectional associations whereby sleep disturbances often precede and are predictive of later depression (Alvaro et al., 2013; Zhai et al., 2015; Zhang et al., 2022). Insomnia sufferers, for example, may be twice as likely to report later depression compared to those without insomnia (Baglioni et al., 2011). Bidirectional associations have been observed from adolescence (Alvaro et al., 2013) to older age (Bao et al., 2017), and within depression sufferers, sleep disturbances are associated with worse depression outcomes, including greater severity, risk of suicidality, and comorbidity with anxiety (Alvaro et al., 2013; Franzen and Buysse, 2008; Wang et al.,

* Corresponding author at: School of Health and Wellbeing, Clarice Pears Building, 90 Byres Road, University of Glasgow, Glasgow G12 8TB, UK.
E-mail address: laura.lyall@glasgow.ac.uk (L.M. Lyall).

https://doi.org/10.1016/j.jad.2023.04.138
Received 28 July 2022; Received in revised form 25 April 2023; Accepted 29 April 2023
Available online 6 May 2023

2019).

Disturbances to circadian rhythms (physiological and behavioural changes that recur over a ~24-hour period) are also associated with depression (Lyall et al., 2018; Walker et al., 2020). Environmental factors causing temporal circadian misalignment, e.g., nightshift work, jetlag, artificial light at night, have been linked to lower mood, and if occurring over a long period, depression onset (Angerer et al., 2017; Bedrosian and Nelson, 2017). Evening chronotype, circadian rhythm sleep disorders or lower amplitude rest-activity rhythms are also associated with MD onset (Byrne et al., 2019; Taylor and Hasler, 2018). Stabilising rhythms, e.g. via light-based chronotherapy, is associated with improvements in depressive symptoms (Perera et al., 2016). As with sleep, a bidirectional relationship with depression is likely, and within depression, circadian disruption is associated with greater severity, with atypical depression, and with suicidality (Carpenter et al., 2021; Courtet and Olié, 2012; Rumble et al., 2020).

Many characteristics of sleep and circadian function have shown association with depression-related outcomes, including sleep efficiency and duration, circadian amplitude and timing. As these are rarely considered together in multivariable models, however, it is unclear if associations reflect intercorrelation between measures or whether specific features of sleep and/or circadian function may be key risk factors. It is also unclear whether objective actigraphy-derived measures provide better predictors of depression and more severe outcomes within depression patients than subjective reports. Greater understanding of which features of sleep and circadian function are most predictive of depression outcomes could improve understanding of mechanisms and inform the development of more targeted interventions.

We examined associations of subjective and objective measures of sleep and circadian rest-activity characteristics with depression-related outcomes in UK Biobank, in penalised regression machine learning models. Outcomes included case vs. control comparisons, i.e., MD cases vs. controls and postnatal depression (PND) vs. controls, and five subdimensions of MD reflecting greater severity (Nguyen et al., 2022): i) atypical vs. typical symptoms; ii) presence vs. absence of comorbid anxiety; iii) severe vs. moderate MD; iv) early vs. later onset; v) MD with vs. without suicidal thoughts/behaviour. Penalised regression was used to identify the most important sleep/circadian predictors of each outcome, and to assess prediction performance.

2. Methods

2.1. Participants

Between 2006 and 2010, over 502,000 UK residents aged 37–73 years were recruited to the UK Biobank. They attended one of 22 assessment centres around the UK and completed sociodemographic, health, lifestyle, mood, cognitive and physical assessments and questionnaires. In 2013–2015, a subset of ~100,000 participants provided up to 7 days of wrist-worn actigraphy data as part of an activity monitoring study (Doherty et al., 2017). In 2016–2017, ~160,000 participants completed an online mental health questionnaire (MHQ). See Fig. 1 in Conroy et al. (2019) for overview of timeline. Here, main analyses are restricted to participants who provided actigraphy data that passed quality control (QC), and for whom sufficient mental health data enabled categorisation into the case or control groups described below (max. n = 64,353). Descriptive sociodemographic and sleep/circadian characteristics for the MD vs. control comparison are provided in Table 1, and sample sizes and descriptive statistics for the other MD dimensions in Supplementary Tables S1–S7. All participants who joined UK Biobank provided written, informed consent, and generic ethical approval was provided by the North West Multi-centre Research Ethics Committee (ref: 21/NW/0157). Analyses were performed using UK Biobank application number 54772 (PI Lyall).

2.2. Predictors

2.2.1. Sociodemographic covariates

Only age, sex, and Townsend deprivation score were included in penalised regression models (alongside sleep/rest-activity variables), but group differences in several other sociodemographic characteristics are summarised in Tables 1 and S1–S7. Data on age (UK Biobank data field #21003), sex (#31), ethnicity (#21000) and educational attainment (#6138) were provided at the baseline assessment. Given relatively small numbers from some ethnic backgrounds, ethnicity in descriptive tables was coded as 'white' and 'non-white', in line with previous publications (Lyall et al., 2018). For education, participants were categorised into those reporting a college/university degree vs. no degree. Townsend deprivation scores (#189) were derived based on postcode of residence: more negative scores reflect greater affluence (Townsend, 1987). Baseline self-report measures were used for smoking status (#20116) and frequency of alcohol intake (#1558). Body-mass index (BMI) was calculated from height and weight measurements at the baseline assessment (weight/(height)^2). During the baseline assessment, participants were asked 12 questions from the Eysenck Personality Questionnaire Revised (Short Form) Neuroticism Scale (Eysenck and Eysenck, 1993), and a composite neuroticism score (range 0–12) was derived (0 if all 12 questions were answered negatively).

2.2.2. Subjective sleep/chronotype characteristics

During the baseline assessment, participants were asked to report on several sleep characteristics. For each, those responding 'do not know' or 'prefer not to answer' were coded as missing. Typical sleep duration was reported as the estimated number of hours sleep in each 24 h, including naps (#1160). We excluded responses below 2 h or above 18 h, and for descriptive tables, responses <7 h were coded as 'short' sleep duration (n = 14,280), 7 or 8 h as 'normal' sleep duration (n = 45,984), and ≥9 h as 'long' sleep (n = 4062), following American Academy of Sleep Medicine guidelines (Ramar et al., 2021). In main analyses, mean subjective sleep duration was included as a continuous variable.

Participants reported how easy they find it to get up in the morning (#1170): we grouped responses into high ('not at all easy'/'not very easy') vs. low ('fairly easy' and 'very easy') difficulty (Sambou et al., 2022). Chronotype (#1180) was coded as definite morning, intermediate (comprising 'more morning than evening person' and 'more evening than morning person' responses) and definite evening. For napping (#1190), insomnia symptoms (#1200; trouble falling asleep at night or waking during the night) and daytime dozing (#1220; unintentionally falling asleep during the day), those responding 'usually' were contrasted with those who responded 'never/rarely' or 'sometimes' (Kyle et al., 2017; Sambou et al., 2022). For snoring (#1210), 'yes' and 'no' responses were contrasted.

2.2.3. Sleep disorders

Participants with a record of a sleep disorder (ICD-10 G47) or non-organic sleep disorder (ICD-10 F51) from linked primary care, hospital admission or death records, and/or self-report (#130921; #131061) were included in a sleep disorder category and compared with those who had no record/report of a sleep disorder.

2.2.4. Objective sleep/rest-activity variables

In 2013–2014, over 100,000 UK Biobank participants (of ~240,000 invited) agreed to take part in an actigraphy-based physical activity monitoring study. They were asked to wear an AX3 triaxial Axivity accelerometer on their dominant wrist for 7 days, while continuing normal activities. Physical activity data were pre-processed by the UK Biobank accelerometer expert working group (Doherty et al., 2017). Among derived measures were the overall acceleration average over the data collection period (milli-gravity units; #90012), standard deviation (SD) of overall acceleration (#90013), and average activity for each hour of the day, across all available days (#90027 - #90050). We also

Table 1
Descriptive sociodemographic and sleep/rest-activity characteristics for broad MD vs. controls.

	Controls			Broad MD			t/χ	p
	N	Mean/N	SD/%	N	Mean/N	SD/%		
Sociodemographic characteristics								
Age	40,124	62.41	7.96	24,229	60.75	7.68	26.01	<0.001
Sex (N, %)	40,124			24,229			2780.48	<0.001
Female		18,714	46.64		16,475	68.00		
Male		21,410	53.36		7754	32.00		
Townsend score	40,090	−1.91	2.72	24,189	−1.47	2.90	−19.27	<0.001
Ethnicity (N, %)	39,985			24,154			46.09	<0.001
Non-white		1464	3.66		646	2.67		
White		38,521	96.34		23,508	97.33		
Education (N, %)	39,946			24,132			20.16	<0.001
No degree		23,110	57.85		13,524	56.04		
Degree		16,836	42.15		10,608	43.96		
BMI	40,045	26.59	4.24	24,170	27.08	4.97	−13.24	<0.001
Neuroticism score	34,098	2.79	2.64	19,984	5.42	3.29	−101.68	<0.001
Smoking status (N, %)	40,006			24,181			351.17	<0.001
Never		23,920	59.79		12,810	52.98		
Former		13,711	34.27		9284	38.39		
Current		2375	5.94		2087	8.63		
Frequency of alcohol consumption (N, %)	40,091			24,217			396.48	<0.001
Never		2063	5.15		1579	6.52		
Occasional		7257	18.10		5768	23.82		
Regular		21,249	53.00		11,667	48.18		
Daily		9522	23.75		5203	21.48		
Season actigraph worn (N, %)	40,124			24,229			12.36	0.01
Spring		8904	22.19		5438	22.44		
Summer		11,076	27.60		6795	28.04		
Autumn		11,246	28.03		6908	28.51		
Winter		8898	22.18		5088	21.00		
Subjective sleep characteristics (baseline)								
Subjective sleep duration (continuous)	40,035	7.18	0.93	24,142	7.15	1.07	4.95	<0.001
Subjective sleep duration (categorised) (N, %)	40,102			24,224			316.06	<0.001
Normal (7-8 h)		29,627	73.88		16,357	67.52		
Short (<7 h)		8267	20.61		6013	24.82		
Long (≥9 h)		2208	5.51		1854	7.65		
Difficulty getting up (N, %)	39,920			24,076			1885.95	<0.001
Easy		35,134	88.01		17,985	74.70		
Not easy		4786	11.99		6091	25.30		
Napping (N, %)	40,092			24,223			12.54	<0.001
No		38,467	95.95		23,100	95.36		
Yes		1625	4.05		1123	4.64		
Insomnia symptoms (N, %)	40,074			24,217			1079.16	<0.001
No		31,254	77.99		16,031	66.20		
Yes		8820	22.01		8186	33.80		
Snoring (N, %)	38,066			22,344			21.28	<0.001
No		23,998	63.04		14,504	64.91		
Yes		14,068	36.96		7840	35.09		
Daytime dozing (N, %)	40,031			24,190			118.41	<0.001
No		39,324	98.23		23,444	96.92		
Yes		707	1.77		746	3.08		
Chronotype (N, %)	35,532			21,938			242.83	<0.001
Morning person		9603	27.03		5323	24.26		
Intermediate		23,203	65.30		14,127	64.40		
Evening person		2726	7.67		2488	11.34		
Health records								
Sleep disorder diagnosis (N, %)	40,124			24,229			435.94	<0.001
No		39,364	98.11		23,070	95.22		
Yes		760	1.89		1159	4.78		
Actigraphy variables (excluding hourly averages)								
Overall acceleration average	40,124	28.29	8.38	24,229	27.54	8.03	11.11	<0.001
Acceleration SD	40,124	56.49	22.06	24,229	53.83	19.44	15.51	<0.001
M10 time	39,959	13.65	1.22	24,119	13.79	1.24	−13.94	<0.001
L5 time	40,012	27.29	1.05	24,149	27.38	1.11	−10.52	<0.001
Sleep midpoint	40,028	26.98	0.91	24,165	27.01	0.95	−4.78	<0.001
Sleep duration (h)	40,028	7.25	0.90	24,165	7.31	0.90	−8.13	<0.001
Sleep duration SD	39,807	0.92	0.57	24,036	0.97	0.58	−10.67	<0.001
Sleep efficiency	40,028	0.76	0.07	24,165	0.76	0.07	−5.24	<0.001
Duration sustained inactivity bouts (daytime)	40,028	0.95	0.66	24,165	1.01	0.72	−10.99	<0.001
Number of nocturnal sleep episodes	40,028	17.25	3.65	24,165	17.30	3.67	−1.74	0.08
Time in bed (h)	40,028	9.67	0.98	24,165	9.71	0.98	−4.47	<0.001

used sleep/rest-activity variables derived by Jones et al. (2019): M10 time - the mean midpoint (in hours past midnight) of the most active 10 h; L5 time - the mean midpoint of the least active 5 h; sleep midpoint time; mean daily night sleep duration; sleep duration SD; mean sleep efficiency (proportion of time in bed spent asleep); mean duration of sustained inactivity bouts during the daytime ('SIBD'); mean number of nocturnal sleep episodes per night (i.e. number of nocturnal awakenings); mean time in bed.

Participants were excluded if any of the following exclusion criteria were met: a) their actigraphy data collection period overlapped with a daylight savings change (#90018); b) they did not have ≥72 hour data, with data in each one-hour period of the day (#90015); c) data was flagged by UK Biobank as not well calibrated (#90016), or as unreliable due to unexpectedly small or large size (#90002), or calibration was not performed using the participant's own data (#90017). Main analyses were restricted to participants with actigraphy data passing the above QC measures (n = 103,670), and who provided sufficient mental health data for categorisation into one or more of the depression/control groups described below (largest sample size = 64,353).

2.3. Depression-related outcome variables

2.3.1. Mental health exclusions from case and control groups

Depression-related outcome variables were largely based on a subset of depression subtypes described by Nguyen et al. (2022). Excluded from all case and control groups were any participants with a diagnosis of schizophrenia, schizoaffective disorder, or bipolar disorder, identified through self-report (#20002), during the MHQ (#20544), or through ICD-10 codes (F20-F29, F30/F31). Participants self-reporting use of antipsychotic medication or lithium (see Nguyen et al.'s Supplementary Table S5 for medication list) were excluded.

2.3.2. Broad major depression (MD) group

A broad MD group consisted of participants who met at least one of the following MD criteria: a) lifetime history of MD or current MD based on MHQ (see Davis et al., 2020); b) ICD-10 code for depressive mood disorder (F32/F33/F34/F38/F39); c) probable MD based on baseline questions, using Smith et al. (2013) criteria; d) self-reported depression during nurse-led interview (#20002) or MHQ (#20544). All participants in depression sub-groups met at least one of the above broad MD criteria.

The control group comprised participants who did not meet any of the above criteria, and additionally: a) did not report having seen a general practitioner (GP; #2090) or psychiatrist (#2100) for 'nerves, anxiety, tension or depression', b) did not report during the MHQ a prolonged (≥2 weeks) period of low mood (#20441) or anhedonia (#20446); c) did not report antidepressant use (see Supplementary Table S5 in Nguyen et al., 2022 for list). For the broad MD comparison, $n_{(MD)} = 24{,}229$; $n_{(control)} = 40{,}124$ (after exclusion of the cases/controls below).

To examine performance of this model for new-onset depression, the above model trained and tested on the broad lifetime history of MD vs. control contrast was also tested on a subset of MD cases (vs. controls) where the first occurrence of a depressive episode was estimated as at least 1 year after all data collection was complete: the 1-year interval was imposed to reduce likelihood of reverse causality (Brunner et al., 2014). There were 321 new-onset MD cases, and a random sample of 530 controls were selected (and removed) from the above broad MD control group (i.e., an equivalent case/control split to the broad MD/control group, ~37 %/63 %) (Shimonovich et al., 2021). The other case-control outcome was PND vs. female controls, and within-case depression outcomes were a) MD cases with vs. without atypical symptoms (based on reversed neurovegetative symptoms: hypersomnia and weight gain during depressive episode; see Brailean et al., 2020); b) MD with vs. without comorbid anxiety disorder; c) severe vs. moderate MD; d) early (≤29 years) vs. later (>40 years) onset MD; e) MD with vs. without suicidality. Inclusion criteria and sample sizes (following actigraphy QC) are described in Table 2.

2.4. Analysis

The following steps were applied to each of the depression outcomes described above. Analyses were conducted in Stata (v16.1); missing data imputation was conducted in R (v4.1.3).

Data for each outcome variable were split into training and test datasets, with a 75/25 split, balancing for age (at time of actigraphy) tertile, sex, season of actigraphy data collection, and the relevant outcome. For training data, continuous variables were standardised to have mean 0 and SD 1: test data were standardised using the mean and SD of the training sample. Separately for training and test data, missing data for predictors were imputed using missForest in R v.4.1.3, an iterative random forest method well-suited to data consisting of both categorical and continuous variables (Stekhoven and Buhlmann, 2012). The number of trees per forest was set to 20.

Penalised regression models were conducted using the imputed data. For comparison, complete cases analyses were conducted and are reported in Supplementary Table S16.

Least absolute shrinkage and selection operator (lasso), ridge and elastic net are penalised regression techniques used to select which of a large number of covariates are useful predictors of the outcome and should be included in the model (Fan et al., 2015; Zou and Hastie, 2005). Regression coefficients are regularised towards zero by selecting a tuning parameter, λ, which determines a penalty term. In lasso, the penalty term, L1-norm, can force coefficients making only a minor contribution to the model to zero, performing variable selection. Ridge regression applies an L2-norm penalty, shrinking coefficients towards zero, but all remain in the resulting model. Elastic net incorporates both lasso and ridge penalty terms, performing variable selection while allowing inclusion of correlated predictors.

Using the training data for each outcome, logistic regression lasso, ridge and elastic net penalised regression models were estimated, in addition to a base logistic regression model. Three basic sociodemographic variables; age, sex (not included in PND models) and Townsend score were forced into each penalised regression model (i.e., their coefficients were not shrunk to zero), alongside 51 sleep/rest-activity predictors (each level of categorical variables is counted as a separate feature). Optimal α values for elastic net, and λ for each of lasso, ridge, and elastic net models, were selected using grid search with nested 10-fold cross-validation (CV) to identify the parameters minimising the CV function: CV was conducted for 128 candidate λ values, and for elastic net, for 9 α values: 0.1–0.9 in increments of 0.1. For lasso, CV-based λ selection was also compared with the adaptive lasso and plug-in methods. Out-of-sample (test data) prediction performance was compared for the three lasso models, elastic net and ridge, and the base logistic regression, and for both penalised and post-selection coefficients. The best-fitting model was selected based on the highest deviance ratio and Receiver Operating Characteristic (ROC) Area Under the Curve (AUC). In the case of equal deviance ratio/AUC, the most parsimonious model with the fewest features was selected.

For each outcome, test data performance is summarised via AUC, accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the point on the ROC curve corresponding to Youden's index (Fluss et al., 2005) (Table 3). Youden's index is calculated from: J = sensitivity + specificity − 1 and corresponds to the maximum height of the ROC curve from the chance line. Youden's index is a common method of finding the optimal trade-off between sensitivity and specificity and therefore represents overall diagnostic performance (Fluss et al., 2005). In a clinical context however, it can be of use to assign greater weight to sensitivity to reduce false negatives (Wu et al., 2021). For comparison, in supplementary material we have also provided summaries of discrimination performance using a weighted Youden's index, where sensitivity is weighted higher (0.55)

Table 2
Description of inclusion criteria and sample size for each case/control group.

Group	Description of inclusion criteria	N
Case-control comparisons		
Broad MD		
Broad MD	Meets at least one of the following criteria: a) Lifetime history of MD or current MD based on MHQ (based on Composite International Diagnostic Interview – Short Form) b) ICD-10 code for depressive mood disorder (F32/F33/F34/F38/F39): UKB first occurrences variables c) probable MD based on baseline mental health questions d) Self-reported depression at baseline interview or during MHQ (Individuals with a first record of MD occurring at least 1 year after actigraphy data collection were excluded from this group.)	24,229
Controls	None of the above MD criteria, and: a) Did not report having seen a general practitioner or psychiatrist for 'nerves, anxiety, tension or depression' b) Did not report during MHQ a period of at least 2 weeks of low mood or anhedonia c) Did not report use of an antidepressant (see ref. Nguyen et al. (2022) for list) Controls randomly selected to form the control group for MD cases occurring onset ≥ 1 year after data collection were also excluded from the broad MD control group.	40,124
MD onset ≥ 1 year after data collection		
Broad MD after data collection	Meets above broad MD criteria, and first occurrence of MD (based on ICD-10 first occurrences variables for F32: #130894; F33: #130896; F34: #130898; F38: #130900; F39: #130902) estimated at least 1 year after collection of all data (i.e., ≥1 year after actigraphy data collection).	321
Controls	A random sample of controls from broad MD control group (then excluded from the former control group), n selected to match broad MD case/control split, i.e., 37%/63%	530
Postnatal depression (PND)		
PND	Among women, those who had given birth (#2734) and met at least one of the following criteria were placed into a PND group: a) self-reported postnatal depression at the baseline interview (#20002) b) had an ICD-10 record of a mental health disorder related to the postpartum period (#130925) c) reported during the MHQ that their worst depressive episode occurred soon after giving birth or was suggested to be postnatal depression (#20445)	1,440
Female controls	Women who had had at least one live birth (#2734) and did not meet broad MD or above PND criteria, did not report help-seeking from a GP/psychiatrist for anxiety or depression and did not report use of antidepressants	15,249
Within-case comparisons		
Atypical vs. typical MD		
Atypical	Above broad MD criteria, and during MHQ reported both hypersomnia (#20534) and weight gain (#20536) during their worst episode of depression.	958
Typical	Broad MD criteria and did not report either hypersomnia or weight gain during worst episode of depression.	18,722
Comorbid anxiety		
MD with comorbid anxiety	Broad MD criteria, and at least one of the following criteria: a) self-reported a diagnosis of anxiety or panic attacks at the baseline nurse-led interview (#20002) b) reported during the MHQ that they had been diagnosed with social anxiety/social phobia, panic attacks, anxiety, nerves, or generalised anxiety disorder (GAD) (#20544) c) ICD-10 record of an anxiety disorder (F40/F41) from linked health records/ self-report (#130905; #130907) d) met criteria for lifetime history of GAD based on MHQ	9,461
MD without comorbid anxiety	Broad MD criteria, but none of the above anxiety disorder criteria.	11,121
Depression severity		
Severe MD	Met criteria for probable MD based on baseline questions (broad MD criterion c above), i.e., reported at least two episodes of feeling anhedonic or depressed for at least 2 weeks. To meet 'severe' criteria, participants additionally reported having seen a psychiatrist for 'nerves, anxiety or depression'.	1,596
Moderate MD	Probable MD criteria from baseline questionnaire as above but reported having seen a GP but not a psychiatrist for 'nerves, anxiety or depression'.	3,131
Early onset		
Early onset	The self-reported age at onset of first episode of depression (#20433 from MHQ) was split into octiles (after removing missing values, and values < 3) for those participants meeting broad MD criteria and with usable actigraphy data. Age at onset within the first three octiles (≤29 years) was defined as 'early onset' depression.	7,583
Later onset	As above, age at onset in the last three octiles was defined as 'later onset' depression.	6,690
Suicidality		
MD with suicidality	Met broad MD criteria, and met at least one of the following criteria: a) self-reported deliberate self-harm or suicide attempt at the baseline interview (#20002) b) Reported that during the 2 weeks preceding the MHQ that they had had suicidal thoughts 'more than half the days' or 'nearly every day' (#20513) c) Reported during the MHQ that they had attempted suicide (#20483)	1289
MD without suicidality	Met broad MD criteria, and during MHQ reported that they had not had recent suicidal thoughts (#20513) and did not self-report deliberate self-harm or suicide attempt at baseline interview (#20002).	18,100

than specificity (0.45) (Li et al., 2013).

3. Results

3.1. Sociodemographic and sleep/circadian characteristics by group

Unadjusted group comparisons for sociodemographic and sleep/rest-activity characteristics (excluding actigraphy hourly averages) are displayed in Table 1 (broad MD vs. control) and Supplementary

Table 3
Summary of discrimination performance for the optimal model (lasso, elastic net, ridge) for each depression-related outcome. Performance metrics (and 95 % confidence intervals) are for the point on the Receiver Operating Characteristic (ROC) curve corresponding to Youden's Index.

Outcome	Optimal model	Youden's Index	AUC	Accuracy	Sensitivity	Specificity	PPV	NPV
Case-control comparisons								
Broad MD vs. controls	Lasso: post-selection	0.26	0.68 (0.67–0.68)	62.85 (62.10–63.60)	64.18 (62.96–65.39)	62.04 (61.08–62.99)	50.60 (49.48–51.73)	74.08 (73.13–75.02)
MD onset after actigraphy	Lasso: post-selection	0.27	–0.70)	68.16 (64.91–71.28)	44.86 (39.33–50.48)	82.26 (78.74–85.42)	60.50 (53.98–66.76)	71.13 (67.36–74.69)
PND	Lasso: penalised	0.24	0.67 (0.64–0.70)	65.82 (64.36–67.26)	57.77 (52.53–62.87)	66.60 (65.07–68.10)	14.30 (12.55–16.18)	94.24 (93.29–95.09)
Within-case comparisons								
Atypical vs. typical MD	Ridge post-selection	0.38	0.74 (0.71–0.77)	67.76 (66.43–69.07)	70.70 (64.12–76.69)	67.63 (66.27–68.97)	9.13 (7.79–10.62)	98.05 (97.51–98.50)
Comorbid anxiety	Ridge: penalised	0.13	0.59 (0.57–0.60)	55.35 (53.98–56.71)	68.56 (66.64–70.43)	44.17 (42.32–46.04)	50.96 (49.21–52.71)	62.41 (60.23–64.55)
Depression severity	Elastic net: post-selection	0.15	0.58 (0.55–0.62)	61.55 (58.71–64.33)	45.06 (40.08–50.12)	69.79 (66.45–72.97)	42.69 (37.89–47.59)	71.78 (68.46–74.94)
Early onset	Ridge: penalised	0.19	0.64 (0.62–0.65)	60.67 (59.05–62.28)	76.12 (74.13–78.03)	43.26 (40.87–45.67)	60.19 (58.20–62.16)	61.65 (58.80–64.44)
Suicidality	Ridge: penalised	0.20	0.63 (0.60–0.66)	68.12 (66.79–69.43)	50.48 (44.81–56.13)	69.35 (67.98–70.69)	10.30 (8.83–11.92)	95.26 (94.48–95.96)

AUC = Area Under the (ROC) Curve; NPV = negative predictive value; PPV = positive predictive value. Optimal model refers to the model with the highest AUC/deviance ratio for each outcome.

Tables S1–S7. Group comparisons for actigraphy-derived hourly activity averages are shown in Supplementary Tables S8–S15.

Participants meeting broad MD criteria were on average younger than controls; had higher Townsend scores, higher BMI and neuroticism scores, a greater proportion were female, of white ethnicity, were more likely to hold a degree, to be a smoker, and were less likely to drink alcohol regularly, consistent with previous UK Biobank studies (Smith et al., 2013).

Participants meeting MD criteria reported fewer hours of sleep per day, but (different subsets of) MD cases were more likely to report either short (<7 h) or long (≥9 h) sleep duration. MD cases more often reported difficulty getting up in the morning (25.30 % vs. 11.99 %), napping and dozing, and were less likely to report snoring. Around 1/3rd of MD participants reported frequent insomnia symptoms, compared to around 1/5th of control participants. A greater percentage of MD cases (4.78 % vs. 1.89 %) had a sleep disorder. MD cases more often reported evening chronotype. For actigraphy-derived variables, MD cases showed lower overall acceleration average, lower variability of acceleration, later M10, L5 and sleep midpoint times, longer sleep duration and greater sleep duration variability, longer duration of daytime inactivity, and longer time in bed compared to controls. MD cases also showed slightly higher sleep efficiency (0.762) compared to controls (0.759), possibly reflecting tendency towards longer overall sleep duration.

For hourly activity averages, MD cases tended to show higher activity in late evening/early night (~9 pm-2 am) and lower activity in early morning to afternoon (~5 am-2 pm) (Table S8).

Findings were similar for the subdivisions of depression, with MD cases meeting criteria for atypical, severe, or early onset depression, or depression with comorbid anxiety or suicidality tending to show greater neuroticism and BMI, and less healthy sleep/circadian characteristics than those with less severe depression, e.g., reduced overall activity, more daytime inactivity and nocturnal awakenings, difficulty getting up, napping, insomnia symptoms, and sleep disorders.

3.2. Penalised regression models

For the comparison of severe vs. moderate MD, a log-likelihood ratio test showed that the base logistic regression model including the sleep/rest-activity predictors did not provide better fit than a model including only age, sex, and Townsend score (LR $\chi^2 = 57.47$, $p = 0.08$) but inclusion of sleep/rest-activity predictors improved the logistic model for all other outcomes ($ps < 0.001$).

For all outcomes, and for both training and test data, the out of bag imputation error estimates using missForest, i.e., normalised root mean squared error (NRMSE) for continuous variables and proportion of falsely classified (PFC) for categorical variables were low, i.e., all $< 8.0 \times 10^{-7}$ for NRMSE and all < 0.15 for PFC.

A summary of optimal model performance is provided in Table 3 (imputed data), and in Supplementary Table S16 for complete cases analyses, using Youden's index to determine the optimal cut-off. Performance when applying a higher weight to sensitivity (0.55) vs. specificity (0.45) using weighted Youden's index is summarised in Table S17 for imputed data and Table S18 for complete cases analysis. ROC curves, alongside a summary of the features selected and their coefficients, ranked in order of magnitude, are summarised in Figs. 1–2 and Supplementary Figs. S1–S5. For binary variables, coefficients for each category were equivalent but with the opposite sign: for ease of interpretation, only the first category listed in model output is referred to in figures.

3.2.1. Broad MD vs. controls

For the broad MD vs. controls comparison, the lasso model (post-selection; $\lambda = 0.001$) provided the highest AUC (0.68; 95 % CI 0.67–0.68), although this was slightly below the 0.70 cut-off for reasonable prediction. Accuracy was 62.85 %, sensitivity 64.18 % and specificity 62.04 % (Table 3). Sensitivity increased to 95.15 % when applying a higher weight to this metric vs. specificity in determining cut-off values, at the cost of lower overall accuracy (45.12 %) and specificity (14.81 %). The lasso model selected 33 predictors: summarised in order of their (post-selection) coefficient magnitude in Fig. 1A–B. Among the top sleep/rest-activity predictors were sleep disorder, difficulty getting up, insomnia symptoms, daytime dozing, napping, actigraphy-derived SIBD, snoring, SD of overall acceleration, evening chronotype and activity from 11 pm–12 am.

The above model was used to generate predicted values for MD cases with onset ≥1 year after data collection. Prediction performance was similar to the broad MD model (AUC = 0.67; 95 % CI 0.60–0.69; accuracy 68.16 %) (Fig. 1C).

3.2.2. Postnatal depression

For the comparison of women having experienced probable PND vs. women who have given birth with no history of depression, the best

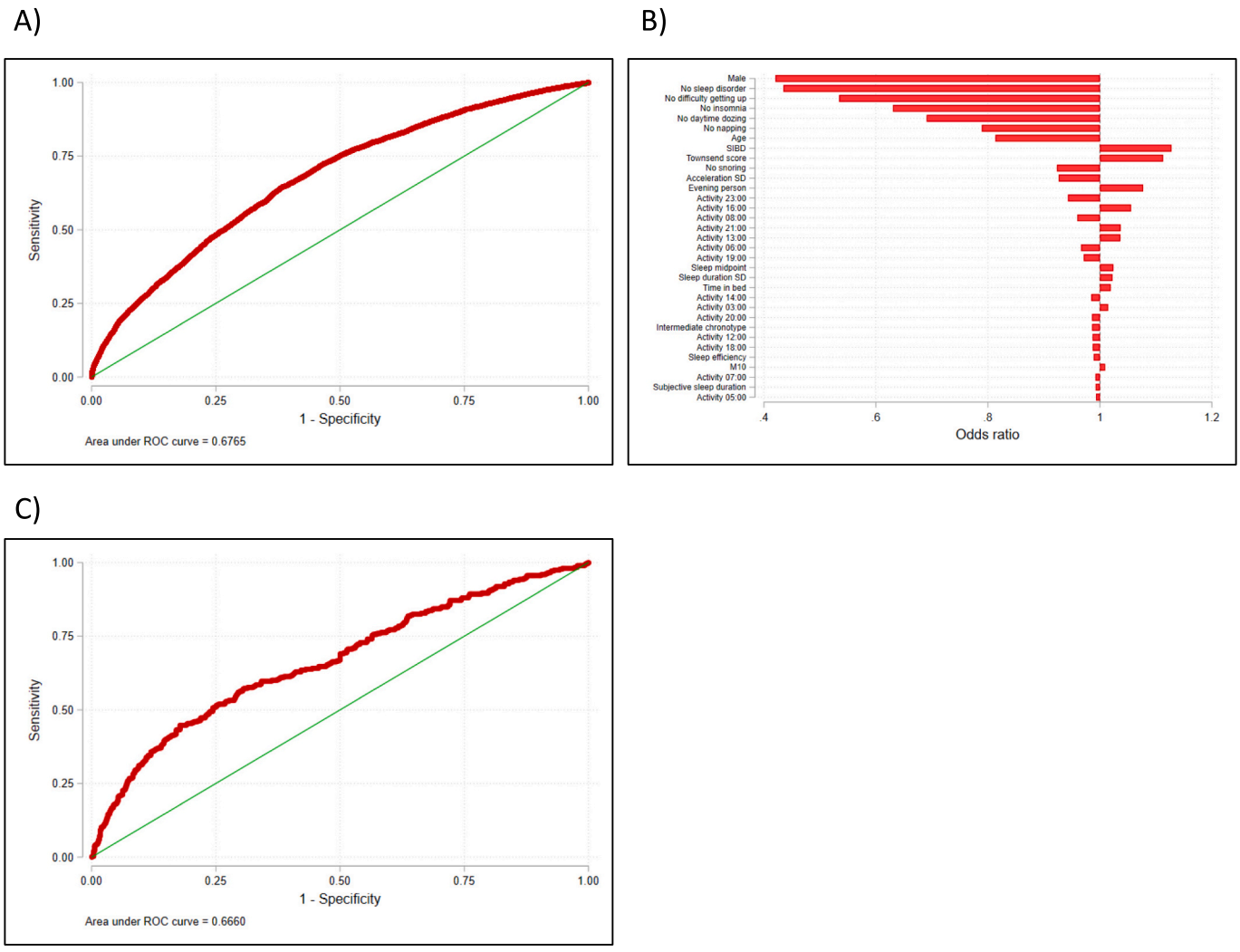

Fig. 1. A) ROC curve for broad MD vs. controls. B) Summary of lasso post-selection coefficients for broad MD vs. controls. C) ROC curve summarising performance of broad MD vs. controls lasso model in separate subset of participants with MD onset ≥ 1 year after data collection.

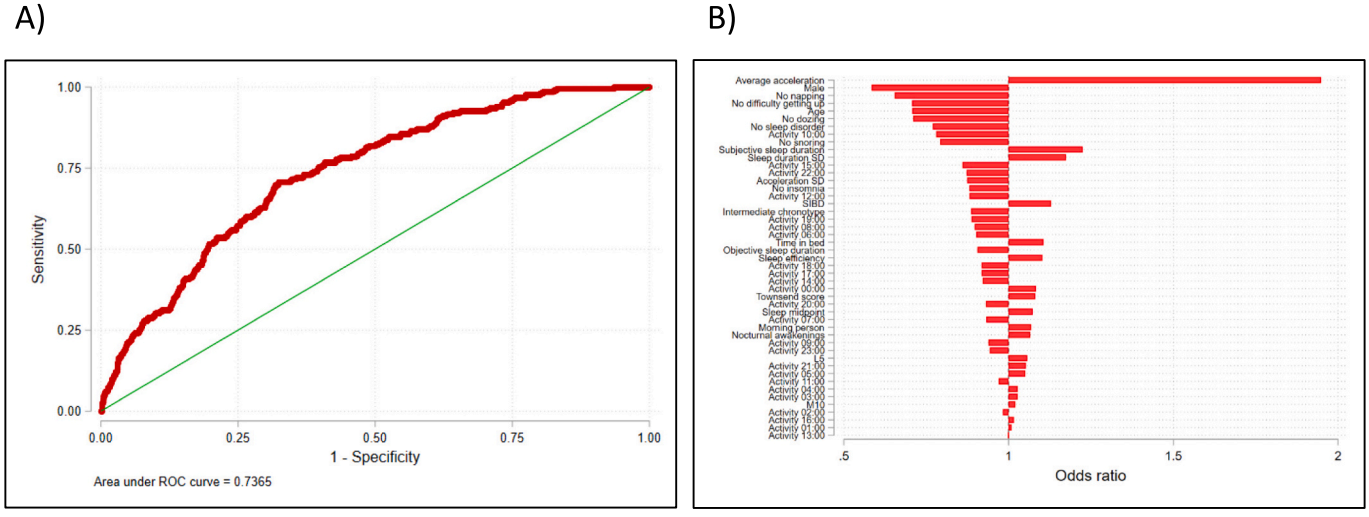

Fig. 2. ROC curve (A) and summary of post-selection ridge regression coefficients (B) for MD with vs. without atypical symptoms.

model was lasso (penalised; $\lambda = 1.5 \times 10^{-3}$; AUC 0.67, CI 0.64–0.70). Twenty-six predictors were selected (Fig. S1).

3.2.3. *Atypical vs. typical MD*

The ridge model with post-selection coefficients ($\lambda = 0.03$) was the best model for MD with vs. without atypical post-selection. The AUC

value of 0.74 (95 % CI 0.71–0.77; Fig. 2) exceeded Hosmer and Lemeshow's (2004) threshold (0.7) for acceptable discrimination. Accuracy, sensitivity, and specificity were all >67 % using Youden's index. When using the weighted Youden's index, accuracy was 25.87 %, sensitivity 97.67 % and specificity 22.57 %.

3.2.4. Comorbid anxiety

Ridge regression (penalised coefficients; $\lambda = 0.07$) provided the highest AUC for depression with vs. without comorbid anxiety, although discrimination was poor (AUC 0.59, 95 % CI 0.57–0.60; see Fig. S2).

3.2.5. Depression severity

An elastic net model (post-selection; $\alpha = 0.9$, $\lambda = 0.02$) yielded the highest AUC for severe vs. moderate depression, but discrimination was poor (AUC 0.58, 95 % CI 0.55–0.62). Five predictors were selected; in order of coefficient size (Fig. S3) these are: sex, difficulty getting up, Townsend score, SIBD and age.

3.2.6. Early onset vs. later onset

Ridge regression provided the highest AUC (0.64, 95 % CI 0.62–0.65) for the comparison of early onset and later onset depression (penalised; $\lambda = 0.17$; Fig. S4).

3.2.7. Suicidality

For the comparison of depression with vs. without suicidal thoughts/behaviours, ridge regression provided the best prediction performance (penalised; $\lambda = 0.03$; AUC 0.63, CI 0.60–0.66, Fig. S5). The top ten predictors in terms of coefficient magnitude were: SIBD, insomnia symptoms, morning activity 7 am–8 am, evening activity 6 pm–7 pm, late evening activity 11 pm-12 am, evening chronotype, morning activity 8 am–9 am, late evening activity 9 pm–10 pm, M10 time and night-time activity 2 am–3 am.

3.3. Most common predictors: association with MD in separate UK Biobank subset

Seven sleep/rest-activity variables were among the 15 largest coefficients in at least half of the final models (imputed and complete cases). These were: difficulty getting up, SIBD, insomnia symptoms, sleep disorder, napping, snoring, and average activity from 8 am–9 am.

Exploratory analyses examined the association of the above subjective/health record markers of poorer sleep (i.e., excluding the actigraphy measures, SIBD and average activity 8-9 am, as participants with these measures were already included in main analyses) with depression outcomes in participants who were not included in main penalised regression analyses as they did not provide usable actigraphy data - either with data but QC was not passed, or actigraphy data was not collected. (For MD broad vs. controls, $n_{(MD)} = 78{,}163$; $n_{(controls)} = 232{,}555$).

A sleep score comprising the number of above (non-actigraphy) factors was calculated: 1 was added to the sleep score for: presence of/frequent: a) difficulty getting up; b) napping, c) insomnia symptoms, d) snoring; e) sleep disorder. As small numbers had the highest score of 5, scores of 4 and 5 were collapsed into a single category. Association of this sleep score with each MD outcome was examined via logistic regression (treating sleep score as categorical). Results are summarised in Table 4.

For each outcome, the OR vs. scores of 0 increased with increasing sleep score: the odds of meeting broad MD criteria were 8.05 times higher for those with at least 4 markers of poor sleep compared to those with a sleep score of 0, and of those with the highest sleep score, 62 % met broad MD criteria. ORs for the highest sleep score vs. 0 for the other outcomes ranged from 1.80 (early onset vs. later onset) to 9.41 (PND vs. controls). Of those meeting MD criteria with sufficient data for categorisation into those with vs. without comorbid anxiety, 73.16 % of those with four or more markers of poor sleep met criteria for comorbid

Table 4
Associations between sleep score and each depression outcome. Sleep score is comprised of the number of (subjective) sleep-related factors shown from the five 'top' predictors identified in most penalised regression models.

Sleep score	MD broad		PND		Atypical		Comorbid anxiety		Severity		Early onset		Suicidality	
	OR	N (%)[a]	OR	N (%)[a]	OR	N (%)[a]	OR	N (%)[a]	OR	N (%)[a]	OR	N (%)[a]	OR	N (%)[a]
0	–	94,497 (80.63)	–	40,342 (98.00)	–	9340 (96.59)	–	9294 (59.83)	–	3439 (67.67)	–	3501 (49.46)	–	9243 (94.42)
1	1.41*	97,755 (76.10)	1.40*	35,186 (97.31)	1.60*	11,129 (94.89)	1.23*	10,934 (55.03)	1.06	4826 (65.76)	1.07	4161 (48.94)	1.21*	10,836 (93.39)
2	2.40*	33,434 (64.97)	2.36*	11,867 (95.59)	2.15*	5639 (93.33)	1.69*	4982 (46.98)	1.27*	2604 (61.29)	1.19*	2055 (46.54)	1.54*	5286 (91.72)
3	4.21*	5954 (51.83)	3.43*	1988 (93.82)	2.98*	1301 (91.11)	2.61*	1077 (36.50)	1.63*	666 (54.37)	1.30*	448 (44.58)	2.58*	1137 (86.86)
4[b]	8.05*	655 (37.60)	9.41*	147 (85.96)	5.71*	195 (84.05)	4.09*	153 (26.84)	2.21*	111 (45.49)	1.80*	58 (36.94)	5.38*	153 (73.56)

	1874 (63.50)		417 (73.16)		133 (54.51)		99 (63.06)		55 (26.44)

Age, sex (excluding PND model) and Townsend score were included in models alongside sleep score.
* Significant (vs. reference category 0) at $p < 0.001$.
[a] For each outcome, the left column N (%) represents controls, the right cases.
[b] Sleep score 4 includes scores of 4 and 5 due to low sample sizes for these scores.

anxiety.

4. Discussion

4.1. Summary

Penalised regression models incorporating the subjective and objective sleep/circadian measures in UK Biobank were constructed to a) examine prediction performance for depression (vs. controls) and more severe outcomes within depression; and b) identify the optimal sleep/circadian predictors of each of the depression-related outcomes. A ridge regression model could discriminate moderately well participants with atypical vs. typical depression (AUC 0.74; 95 % CI 0.71–0.77). Discrimination was relatively poor for other models (AUCs < 0.70), although for MD vs. controls (including a separate subset of participants with onset of depression $\geq$1 year after data collection), and PND vs. controls, lasso models were close to the threshold for acceptable discrimination (MD: AUC 0.68, 95 % CI 0.67–0.68; PND: AUC 0.67; CI 0.64–0.70).

Several potentially important clinical markers of depression susceptibility were identified via feature selection and/or relative coefficient magnitude: insomnia symptoms, snoring, napping, difficulty getting up, sleep disorder, and actigraphy-derived duration of daytime inactivity (SIBD) and lower morning activity around 8 am. In follow-up analyses within a separate subset of UK Biobank participants, individuals self-reporting any of the first five (excluding actigraphy measures) were more likely to experience depression or more severe depression outcomes, particularly if a combination of these markers was present.

4.2. Discrimination performance

Sleep problems and circadian disruption are common in the general population, including in those who are otherwise healthy or suffer from psychiatric disorders other than depression (Walker et al., 2020). It is perhaps unsurprising therefore that sleep and circadian measures alone (alongside limited sociodemographic characteristics) did not discriminate those suffering from depression or more severe depression with a high degree of accuracy. Of note, a more liberal threshold for 'acceptable' discrimination of AUC > 0.6 is often applied (Cha et al., 2022; Yang and Berdine, 2017): all but two of our models surpassed this threshold, but only one exceeded the more conservative threshold applied here of AUC > 0.7 (Hosmer and Lemeshow, 2004). The fact that inclusion of sleep/rest-activity predictors improved model performance beyond basic sociodemographics (for all outcomes except severe vs. moderate MD) and that the selected features were strongly associated with all outcomes reinforces the importance of the selected sleep/circadian variables in depression, and outcomes in depression. Consistent with this, sleep and circadian disruption have been linked longitudinally to onset of and worse outcomes in depression (Zhang et al., 2022) and may act as general markers of vulnerability to mental ill-health (Lyall et al., 2018).

The highest discrimination was for depression with atypical vs. typical symptoms. As the definition of lifetime history of atypical depression included (self-reported) hypersomnia during the worst period of depression (Brailean et al., 2020; Nguyen et al., 2022), it is perhaps unsurprising that detailed sleep/rest-activity characteristics were useful in discriminating these two groups. However, it should be noted that i) 94 % of the atypical depression group did not meet criteria for current depression at the time of the MHQ, and ii) all other contrasts (MD/PND vs. controls and each depression subtype) were not defined based on any sleep characteristics but also showed widespread sleep/rest-activity differences. Findings are therefore consistent with suggestions that both depression itself, and more severe depression subtypes show reliable association with sleep and rest-activity characteristics regardless of current depressive state.

Discrimination performance may improve with addition of more fine-grained actigraphy data, e.g., epoched activity counts across the duration of the data collection period. Other machine learning methods, particularly non-linear approaches e.g., Support Vector Machines (SVM) with radial basis function or gradient boosting machine could also improve performance. Here, however, a principal aim was feature selection and interpretability of results, which penalised regression is well-suited to, and for which it has shown good performance in clinical psychiatric prediction models (Chekroud et al., 2021). Initial scoping using SVM-recursive feature elimination with several outcomes did not improve performance, consistent with a systematic review finding no improvement in clinical prediction for non-linear machine learning methods compared to logistic regression including penalised regression (Christodoulou et al., 2019). As the number of predictors here was relatively low, alternative non-linear machine learning methods may have a greater influence on performance in future, higher-dimensional studies where larger numbers of lifestyle, neuroimaging and genetic features are included, and non-linear effects are more likely.

Mental health data were not available at the time of actigraphy data collection, but at the time of the baseline and online follow-up MHQs, only around 2 % of participants with lifetime history of depression were thought to be currently experiencing a depressive episode. Findings therefore support suggestions that individuals with depression differ in terms of sleep/rest-activity characteristics from healthy controls even during periods of euthymia, and mirror previous findings of strong, reliable association between general sleep/rest-activity features and lifetime history of psychiatric disorders including depression (Lyall et al., 2018; Wainberg et al., 2021). Although addressing a different question, it is likely that classification accuracy would be improved if analyses were restricted to participants currently suffering a depressive episode.

It should be noted that in a clinical context, simple screening tools such as the Patient Health Questionnaire (including PHQ-9 and PHQ-2) outperform the above models in terms of AUC, sensitivity and specificity, and therefore remain the most efficient option for rapid screening of both depression diagnosis and depression severity (Cameron et al., 2011; Gilbody et al., 2007; Manea et al., 2016). Here, applying greater weight to sensitivity (Tables S17-S18) improved ability to detect cases, maximising sensitivity and NPV, at the cost of lowered specificity, PPV and overall accuracy. It could be argued that such weighting could be useful in a clinical context to capture all patients who may benefit from further screening and/or intervention, to reduce false negatives (Chubak et al., 2012; Wu et al., 2021). Our findings contribute to suggestions that individuals suffering from disturbed sleep and rest-activity patterns may be more likely to suffer from depression and more severe outcomes within depression regardless of current mood state, and may benefit from greater monitoring (Lovato and Gradisar, 2014; Wang et al., 2019).

4.3. Optimal predictors

A key aim was to identify which of the many available sleep and rest-activity variables were the most important predictors of depression-related outcomes. Several features occurred in the top 15 predictors in at least half of the models: difficulty getting up, insomnia symptoms, snoring, napping, sleep disorder, and actigraphy-derived SIBD and average activity around 8 am. The non-actigraphy measures from this list were strongly associated with the depression outcomes in a separate subset of UK Biobank participants who had not provided usable actigraphy data, particularly where multiple factors occurred in combination.

This set of optimal predictors may be consistent with disturbed sleep during the night (insomnia, snoring) resulting in sleepiness and inactivity during the day (difficulty getting up, napping, SIBD, lower morning activity) (Berger et al., 2021) as contributing factors to depression occurrence and greater severity. This may implicate insufficient/disturbed sleep as a greater contributor to depression-related

outcomes than circadian disruption per se. Other circadian factors such as chronotype, and timing of actigraphy-derived M10, L5 and sleep midpoint were not among the top predictors when accounting for the full range of sleep/activity measures, despite previous findings linking later chronotype to depression risk (Vetter et al., 2018). Lack of circadian features among the top predictors may however have been linked to their collinearity with morning activity.

Two of the top predictors were based on actigraphy measurements: SIBD (duration of sustained inactivity bouts during the daytime), and activity between 8 am and 9 am. Coefficients and descriptive tables indicate those who spend more time inactive during the day and are less active in the morning are more likely to suffer from depression, and more severe depression, including depression with suicidality. This is consistent with recent findings that activity deficits are most pronounced in the morning in adults with depression symptoms (Smagula et al., 2021). It has been proposed that this association of reduced morning activity with depressive symptoms may be linked to sociopsychological factors such as reduced engagement in social and physical activity leading to rumination (Smagula et al., 2021). Reduced circadian entrainment and morning daylight exposure may be another factor: this has been linked to delays in sleep timing, shorter sleep duration, and reduced sleep quality (Blume et al., 2019). Some evidence suggests specific targeting of morning activity could be a useful clinical intervention for depression (Smagula et al., 2022). Findings that greater morning and evening activity and lower nocturnal activity were linked to reduced odds of suicidality are consistent with previous literature demonstrating association of reduced physical activity and increased sedentary behaviour with suicide risk (Vancampfort et al., 2019, 2018): targeting of activity levels has been proposed as an intervention, and it is possible that activity monitoring could be of use in identifying individuals at greater risk of serious outcomes such as suicidality (Vancampfort et al., 2018).

While actigraphy-measured inactivity during the day, particularly morning-time, appears to be an important predictor of depression-related outcomes, most of the top predictors were based on self-report. Detailed actigraphy assessment over days/weeks may not be required to determine those more likely to show depression and more severe outcomes within depression, particularly as our exploratory analyses excluding actigraph-based measures demonstrated strong association of the subjective measures with depression and its severity. SIBD and morning activity could potentially be incorporated into questionnaires alongside the above self-reported sleep characteristics, for example at GP/psychiatrist assessments of those with depression, to identify individuals at greater risk of worse outcomes. While self-report can be flawed, e.g. sedentary behaviour is often underestimated and sleep problems overestimated, subjective measures show reasonable reliability (Alfano et al., 2015; Prince et al., 2020), and are less costly and invasive compared to actigraphy. Replication of current findings in other cohorts such as Adolescent Brain Cognitive Development (ABCD) study may help clarify a) whether the factors identified here are useful markers of vulnerability to depressive disorders; and b) if so, whether subjective self-report measures are sufficient to identify at-risk groups.

4.4. Limitations

The intervals between collection of different data types may have limited discrimination performance: subjective sleep measures were collected during the baseline assessment (2006–2010) alongside some mental health data, actigraphy data were collected between 2013 and 2015, Mental Health Questionnaire data from 2016 to 2017, and linked health records could come from any time point before/after recruitment. Although sleep disturbances are relatively stable over time (Fatima et al., 2020), participants' sleep/activity habits may have varied between data collection periods, and/or may not have been representative of their typical habits or those around the time of the relevant depression outcome.

Similarly, as UK Biobank is a cohort of middle- to older-age adults, participants' sleep/circadian characteristics at the time of data collection may not reflect those from around the time of their depression onset (typically during adolescence/early adulthood; Solmi et al., 2022) potentially limiting associations and discrimination performance. Similar analyses in younger cohorts, e.g., ABCD, may yield greater insights into key risk factors as well as testing generalisability of the current prediction models. This is particularly important as actigraphy devices and methodology can differ markedly between studies and samples, further highlighting the need for assessment of generalisability (Fekedulegn et al., 2020; Patterson et al., 2023).

The restriction of analyses to those with valid actigraphy data allowed inclusion of all available sleep/rest-activity measures, both subjective and objective. Among this subsample, almost 40 % met criteria for lifetime history of depression: this is higher than previous estimates from the full UK Biobank sample or from other population studies (~27 %; Smith et al., 2013). This is likely due to exacerbated selection bias within the subsample, whereby individuals opting into the follow-up actigraphy investigation were slightly more likely to suffer from depression (see Lyall et al., 2022 for similar findings in the neuroimaging subsample). It is unclear why this might be, particularly as follow-up subsamples are typically associated with 'healthy bias' (Fry et al., 2017), but could be linked to increased interest in participating in health research among those suffering from depression (Adams et al., 2020). Reassuringly however, in exploratory analyses in an independent subset of UK Biobank participants without available actigraphy data (with lower depression prevalence), the key subjective sleep variables identified in most models were strongly associated with the depression-related outcomes. The UK Biobank's use of a single protocol and large sample size mean that while the sample and particularly subsamples may not be representative of the general population and should not be used for prevalence estimates, findings relating to exposure-disease associations are likely to be generalisable (Fry et al., 2017).

The direction of causality (if any) between sleep/circadian disruption and depression-related outcomes is not of current relevance to the prediction models reported: the predictors could contribute to discrimination whether they are a cause or consequence of the outcome measures, or both are linked to other, unmeasured factors. Causality is however of great theoretical and clinical interest, and future studies e.g., involving Mendelian Randomisation, and causal modelling including in younger cohorts (e.g., ABCD) will aim to address this.

5. Conclusions

Penalised regression models incorporating sleep and circadian rest-activity characteristics were able to discriminate moderately well those suffering depression with vs. without atypical symptoms, and discrimination of MD vs. controls and PND vs. female controls approached the cut-off for reasonable discrimination. Prediction for other depression-related outcomes (including MD with vs. without comorbid anxiety) was poor. Findings highlight several potentially important sleep/rest-activity related predictors of depression and its severity. Individuals experiencing difficulty getting up in the morning, insomnia symptoms, greater inactivity during the daytime, lower morning activity, sleep disorders, and who snore or take naps may be more likely to have a lifetime history of depression and worse depression outcomes, particularly when several of these factors co-occur together. Future studies incorporating these factors alongside additional sociodemographic, genetic and neuroimaging data may lead to improved discrimination performance.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jad.2023.04.138.

Role of the funding source

This research was conducted using the UK Biobank resource, under

application 54772 (PI Lyall). UK Biobank was established by the Wellcome Trust, Medical Research Council, Department of Health, Scottish Government and Northwest Regional Development Agency. UK Biobank has also had funding from the Welsh Assembly Government and the British Heart Foundation. Data collection was funded by UK Biobank. LML is supported by a JMAS Sim Fellowship for depression research from the Royal College of Physicians of Edinburgh, and a Lord Kelvin Adam Smith (LKAS) Fellowship from the University of Glasgow. RJS is supported by a University of Glasgow LKAS fellowship and a UKRI Innovation-**HDR-UK** Fellowship (MR/S003061/1). DJS acknowledges the support of a Lister Prize Fellowship (173096). NS is funded by a Medical Research Council Doctoral Training Programme in Precision Medicine PhD studentship. The funders had no role in the design or analysis of this study, decision to publish, or preparation of the manuscript.

CRediT authorship contribution statement

All authors contributed to the design of the study. LML conducted data analysis and drafted the manuscript. All authors contributed to editing of the manuscript and have approved the final version of the manuscript.

Declaration of competing interest

None.

Acknowledgements

The authors are grateful to the participants of UK Biobank.

References

Adams, M.J., Hill, W.D., Howard, D.M., Dashti, H.S., Davis, K.A.S., Campbell, A., Clarke, T.-K., Deary, I.J., Hayward, C., Porteous, D., Hotopf, M., McIntosh, A.M., 2020. Factors associated with sharing e-mail information and mental health survey participation in large population cohorts. Int. J. Epidemiol. 49, 410–421. https://doi.org/10.1093/ije/dyz134.

Alfano, C.A., Patriquin, M.A., De Los Reyes, A., 2015. Subjective – objective sleep comparisons and discrepancies among clinically-anxious and healthy children. J. Abnorm. Child Psychol. 43, 1343–1353. https://doi.org/10.1007/s10802-015-0018-7.

Alvaro, P.K., Roberts, R.M., Harris, J.K., 2013. A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. Sleep 36, 1059–1068. https://doi.org/10.5665/sleep.2810.

Angerer, P., Schmook, R., Elfantel, I., Li, J., 2017. Night work and the risk of depression. Dtsch. Arztebl. Int. 114, 404–411. https://doi.org/10.3238/arztebl.2017.0404.

Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., Lombardo, C., Riemann, D., 2011. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. J. Affect. Disord. 135, 10–19. https://doi.org/10.1016/j.jad.2011.01.011.

Bao, Y.-P., Han, Y., Ma, J., Wang, R.-J., Shi, L., Wang, T.-Y., He, J., Yue, J.-L., Shi, J., Tang, X.-D., Lu, L., 2017. Cooccurrence and bidirectional prediction of sleep disturbances and depression in older adults: meta-analysis and systematic review. Neurosci. Biobehav. Rev. 75, 257–273. https://doi.org/10.1016/j.neubiorev.2017.01.032.

Bedrosian, T.A., Nelson, R.J., 2017. Timing of light exposure affects mood and brain circuits. Transl. Psychiatry 7. https://doi.org/10.1038/tp.2016.262 e1017-e1017.

Berger, M., Hirotsu, C., Haba-Rubio, J., Betta, M., Bernardi, G., Siclari, F., Waeber, G., Vollenweider, P., Marques-Vidal, P., Heinzer, R., 2021. Risk factors of excessive daytime sleepiness in a prospective population-based cohort. J. Sleep Res. 30, e13069 https://doi.org/10.1111/jsr.13069.

Blume, C., Garbazza, C., Spitschan, M., 2019. Effects of light on human circadian rhythms, sleep and mood. Somnologie 23, 147. https://doi.org/10.1007/s11818-019-00215-x.

Brailean, A., Curtis, J., Davis, K., Dregan, A., Hotopf, M., 2020. Characteristics, comorbidities, and correlates of atypical depression: evidence from the UK biobank mental health survey. Psychol. Med. 50, 1129–1138. https://doi.org/10.1017/S0033291719001004.

Brunner, E.J., Shipley, M.J., Britton, A.R., Stansfeld, S.A., Heuschmann, P.U., Rudd, A.G., Wolfe, C.D., Singh-Manoux, A., Kivimaki, M., 2014. Depressive disorder, coronary heart disease, and stroke: dose–response and reverse causation effects in the Whitehall II cohort study. Eur. J. Prev. Cardiol. 21, 340–346. https://doi.org/10.1177/2047487314520785.

Byrne, E.M., Timmerman, A., Wray, N.R., Agerbo, E., 2019. Sleep disorders and risk of incident depression: a population case-control study. Twin Res. Hum. Genet. 22, 140–146. https://doi.org/10.1017/thg.2019.22.

Cameron, I.M., Cardy, A., Crawford, J.R., du Toit, S.W., Hay, S., Lawton, K., Mitchell, K., Sharma, S., Shivaprasad, S., Winning, S., Reid, I.C., 2011. Measuring depression severity in general practice: discriminatory performance of the PHQ-9, HADS-D, and BDI-II. Br. J. Gen. Pract. 61, e419–e426. https://doi.org/10.3399/bjgp11X583209.

Carpenter, J.S., Crouse, J.J., Scott, E.M., Naismith, S.L., Wilson, C., Scott, J., Merikangas, K.R., Hickie, I.B., 2021. Circadian depression: a mood disorder phenotype. Neurosci. Biobehav. Rev. 126, 79–101. https://doi.org/10.1016/j.neubiorev.2021.02.045.

Cha, J., Spielberg, J.M., Hu, B., Altinay, M., Anand, A., 2022. Differences in network properties of the structural connectome in bipolar and unipolar depression. Psychiatry Res. Neuroimaging 321, 111442. https://doi.org/10.1016/j.pscychresns.2022.111442.

Chekroud, A.M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., Cohen, Z., Belgrave, D., DeRubeis, R., Iniesta, R., Dwyer, D., Choi, K., 2021. The promise of machine learning in predicting treatment outcomes in psychiatry. World Psychiatry 20, 154–170. https://doi.org/10.1002/wps.20882.

Christodoulou, E., Ma, J., Collins, G.S., Steyerberg, E.W., Verbakel, J.Y., Van Calster, B., 2019. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J. Clin. Epidemiol. 110, 12–22. https://doi.org/10.1016/j.jclinepi.2019.02.004.

Chubak, J., Pocobelli, G., Weiss, N.S., 2012. Trade-offs between accuracy measures for electronic healthcare data algorithms. J. Clin. Epidemiol. 65, 343–349. https://doi.org/10.1016/j.jclinepi.2011.09.002 e2.

Conroy, M., Sellors, J., Effingham, M., Littlejohns, T.J., Boultwood, C., Gillions, L., Sudlow, C.L.M., Collins, R., Allen, N.E., 2019. The advantages of UK Biobank's open-access strategy for health research. J. Intern. Med. 286, 389–397. https://doi.org/10.1111/joim.12955.

Courtet, P., Olié, E., 2012. Circadian dimension and severity of depression. Eur. Neuropsychopharmacol. 22, S476–S481. https://doi.org/10.1016/j.euroneuro.2012.07.009.

Davis, K.A.S., Coleman, J.R.I., Adams, M., Allen, N., Breen, G., Cullen, B., Dickens, C., Fox, E., Graham, N., Holliday, J., Howard, L.M., John, A., Lee, W., McCabe, R., McIntosh, A., Pearsall, R., Smith, D.J., Sudlow, C., Ward, J., Zammit, S., Hotopf, M., 2020. Mental health in UK biobank – development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. BJPsych Open 6. https://doi.org/10.1192/bjo.2019.100.

Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M.H., White, T., van Hees, V.T., Trenell, M.I., Owen, C.G., Preece, S.J., Gillions, R., Sheard, S., Peakman, T., Brage, S., Wareham, N.J., 2017. Large scale population assessment of physical activity using wrist worn accelerometers: the UK biobank study. PLOS ONE 12, e0169649. https://doi.org/10.1371/journal.pone.0169649.

Eysenck, H.J., Eysenck, S.B.G., 1993. Eysenck Personality Questionnaire-Revised. https://doi.org/10.1037/t05461-000.

Fan, L., Chen, S., Li, Q., Zhu, Z., 2015. Variable selection and model prediction based on Lasso, adaptive lasso and elastic net. In: 2015 4th International Conference on Computer Science and Network Technology (ICCSNT). Presented at the 2015 4th International Conference on Computer Science and Network Technology (ICCSNT), pp. 579–583. https://doi.org/10.1109/ICCSNT.2015.7490813.

Fatima, Y., Bucks, R.S., Mamun, A.A., Skinner, I., Rosenzweig, I., Leschziner, G., Skinner, T.C., 2020. Sleep trajectories and mediators of poor sleep: findings from the longitudinal analysis of 41,094 participants of the UK biobank cohort. Sleep Med. 76, 120–127. https://doi.org/10.1016/j.sleep.2020.10.020.

Fekedulegn, D., Andrew, M.E., Shi, M., Violanti, J.M., Knox, S., Innes, K.E., 2020. Actigraphy-based assessment of sleep parameters. Ann. Work Expo. Health 64, 350–367. https://doi.org/10.1093/annweh/wxaa007.

Fluss, R., Faraggi, D., Reiser, B., 2005. Estimation of the youden index and its associated cutoff point. Biom. J. 47, 458–472. https://doi.org/10.1002/bimj.200410135.

Franzen, P.L., Buysse, D.J., 2008. Sleep disturbances and depression: risk relationships for subsequent depression and therapeutic implications. Dialogues Clin. Neurosci. 10, 473–481. https://doi.org/10.31887/DCNS.2008.10.4/plfranzen.

Fry, A., Littlejohns, T.J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., Allen, N.E., 2017. Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. Am. J. Epidemiol. 186, 1026–1034. https://doi.org/10.1093/aje/kwx246.

Gilbody, S., Richards, D., Brealey, S., Hewitt, C., 2007. Screening for depression in medical settings with the patient health questionnaire (PHQ): a diagnostic meta-analysis. J. Gen. Intern. Med. 22, 1596–1602. https://doi.org/10.1007/s11606-007-0333-y.

Hosmer, D.W.J., Lemeshow, S., 2004. Applied logistic regression [WWW document]. CERN document server. URL. https://cds.cern.ch/record/997237. (Accessed 23 May 2022).

Johnston, K.M., Powell, L.C., Anderson, I.M., Szabo, S., Cline, S., 2019. The burden of treatment-resistant depression: a systematic review of the economic and quality of life literature. J. Affect. Disord. 242, 195–210. https://doi.org/10.1016/j.jad.2018.06.045.

Jones, S.E., van Hees, V.T., Mazzotti, D.R., Marques-Vidal, P., Sabia, S., van der Spek, A., Dashti, H.S., Engmann, J., Kocevska, D., Tyrrell, J., Beaumont, R.N., Hillsdon, M., Ruth, K.S., Tuke, M.A., Yaghootkar, H., Sharp, S.A., Ji, Y., Harrison, J.W., Freathy, R.M., Murray, A., Luik, A.I., Amin, N., Lane, J.M., Saxena, R., Rutter, M.K., Tiemeier, H., Kutalik, Z., Kumari, M., Frayling, T.M., Weedon, M.N., Gehrman, P.R., Wood, A.R., 2019. Genetic studies of accelerometer-based sleep measures yield new insights into human sleep behaviour. Nat. Commun. 10, 1585. https://doi.org/10.1038/s41467-019-09576-1.

Kyle, S.D., Sexton, C.E., Feige, B., Luik, A.I., Lane, J., Saxena, R., Anderson, S.G., Bechtold, D.A., Dixon, W., Little, M.A., Ray, D., Riemann, D., Espie, C.A., Rutter, M. K., Spiegelhalder, K., 2017. Sleep and cognitive performance: cross-sectional associations in the UK biobank. Sleep Med. 38, 85–91. https://doi.org/10.1016/j.sleep.2017.07.001.

Li, D., Shen, F., Yin, Y., Peng, J., Chen, P., 2013. Weighted youden index and its two-independent-sample comparison based on weighted sensitivity and specificity. Chin. Med. J. 126, 1150. https://doi.org/10.3760/cma.j.issn.0366-6999.20123102.

Lovato, N., Gradisar, M., 2014. A meta-analysis and model of the relationship between sleep and depression in adolescents: recommendations for future research and clinical practice. Sleep Med. Rev. 18, 521–529. https://doi.org/10.1016/j.smrv.2014.03.006.

Lyall, L.M., Wyse, C.A., Graham, N., Ferguson, A., Lyall, D.M., Cullen, B., Celis Morales, C.A., Biello, S.M., Mackay, D., Ward, J., Strawbridge, R.J., Gill, J.M.R., Bailey, M.E.S., Pell, J.P., Smith, D.J., 2018. Association of disrupted circadian rhythmicity with mood disorders, subjective wellbeing, and cognitive function: a cross-sectional study of 91 105 participants from the UK biobank. Lancet Psychiatry 5, 507–514. https://doi.org/10.1016/S2215-0366(18)30139-1.

Lyall, D.M., Quinn, T., Lyall, L.M., Ward, J., Anderson, J.J., Smith, D.J., Stewart, W., Strawbridge, R.J., Bailey, M.E.S., Cullen, D., 2022. Quantifying bias in psychological and physical health in the UK Biobank imaging sub-sample. Brain Commun. 4 (3), fcac119 https://doi.org/10.1093/braincomms/fcac119.

Manea, L., Gilbody, S., Hewitt, C., North, A., Plummer, F., Richardson, R., Thombs, B.D., Williams, B., McMillan, D., 2016. Identifying depression with the PHQ-2: a diagnostic meta-analysis. J. Affect. Disord. 203, 382–395. https://doi.org/10.1016/j.jad.2016.06.003.

Nguyen, T.-D., Harder, A., Xiong, Y., Kowalec, K., Hägg, S., Cai, N., Kuja-Halkola, R., Dalman, C., Sullivan, P.F., Lu, Y., 2022. Genetic heterogeneity and subtypes of major depression. Mol. Psychiatry 1–9. https://doi.org/10.1038/s41380-021-01413-6.

Patterson, M.R., Nunes, A.A.S., Gerstel, D., Pilkar, R., Guthrie, T., Neishabouri, A., Guo, C.C., 2023. 40 years of actigraphy in sleep medicine and current state of the art algorithms. npj Digit. Med. 6, 1–7. https://doi.org/10.1038/s41746-023-00802-1.

Perera, S., Eisen, R., Bhatt, M., Bhatnagar, N., de Souza, R., Thabane, L., Samaan, Z., 2016. Light therapy for non-seasonal depression: systematic review and meta-analysis. BJPsych Open 2, 116–126. https://doi.org/10.1192/bjpo.bp.115.001610.

Prince, S.A., Cardilli, L., Reed, J.L., Saunders, T.J., Kite, C., Douillette, K., Fournier, K., Buckley, J.P., 2020. A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. Int. J. Behav. Nutr. Phys. Act. 17, 31. https://doi.org/10.1186/s12966-020-00938-3.

Ramar, K., Malhotra, R.K., Carden, K.A., Martin, J.L., Abbasi-Feinberg, F., Aurora, R.N., Kapur, V.K., Olson, E.J., Rosen, C.L., Rowley, J.A., Shelgikar, A.V., Trotti, L.M., 2021. Sleep is essential to health: an American Academy of sleep medicine position statement. J. Clin. Sleep Med. 17, 2115–2119. https://doi.org/10.5664/jcsm.9476.

Riemann, D., Berger, M., Voderholzer, U., 2001. Sleep and depression — results from psychobiological studies: an overview. Biol. Psychol. 57, 67–103. https://doi.org/10.1016/S0301-0511(01)00090-4.

Rumble, M.E., McCall, W.V., Dickson, D.A., Krystal, A.D., Rosenquist, P.B., Benca, R.M., 2020. An exploratory analysis of the association of circadian rhythm dysregulation and insomnia with suicidal ideation over the course of treatment in individuals with depression, insomnia, and suicidal ideation. J. Clin. Sleep Med. 16, 1311–1319. https://doi.org/10.5664/jcsm.8508.

Sambou, M.L., Zhao, X., Hong, T., Naveed, M., Sambou, A., El Hafa, F., Basnet, T.I.B., Dai, J., 2022. Investigation of the relationships between sleep behaviors and risk of healthspan termination: a prospective cohort study based on 323,373 UK-Biobank participants. Sleep Breath. 26, 205–213. https://doi.org/10.1007/s11325-021-02394-0.

Shimonovich, M., Pearce, A., Thomson, H., Keyes, K., Katikireddi, S.V., 2021. Assessing causality in epidemiology: revisiting Bradford Hill to incorporate developments in causal thinking. Eur. J. Epidemiol. 36, 873–887. https://doi.org/10.1007/s10654-020-00703-7.

Smagula, S.F., Capps, C.S., Krafty, R.T., 2021. Evaluating the timing of differences in activity related to depression symptoms across adulthood in the United States. J. Affect. Disord. 284, 64–68. https://doi.org/10.1016/j.jad.2021.01.069.

Smagula, S.F., Isenberg, A.S., Stahl, S.T., Alexopoulos, G.S., Reynolds, C.F., Rodakowski, J., 2022. Adaptation and pilot study of a behavioral intervention targeting morning activation deficits in dementia caregivers: scheduling activity and monitoring mornings (SAMM). Am. J. Geriatr. Psychiatry 30, 494–503. https://doi.org/10.1016/j.jagp.2021.09.017.

Smith, D.J., Nicholl, B.I., Cullen, B., Martin, D., Ul-Haq, Z., Evans, J., Gill, J.M.R., Roberts, B., Gallacher, J., Mackay, D., Hotopf, M., Deary, I., Craddock, N., Pell, J.P., 2013. Prevalence and characteristics of probable major depression and bipolar disorder within UK biobank: cross-sectional study of 172,751 participants. PLOS ONE 8, e75362. https://doi.org/10.1371/journal.pone.0075362.

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar de Pablo, G., Il Shin, J., Kirkbride, J.B., Jones, P., Kim, J.H., Kim, J.Y., Carvalho, A.F., Seeman, M.V., Correll, C.U., Fusar-Poli, P., 2022. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. Mol. Psychiatry 27, 281–295. https://doi.org/10.1038/s41380-021-01161-7.

Stekhoven, D.J., Buhlmann, P., 2012. MissForest–non-parametric missing value imputation for mixed-type data. Bioinformatics 28, 112–118. https://doi.org/10.1093/bioinformatics/btr597.

Taylor, B.J., Hasler, B.P., 2018. Chronotype and mental health: recent advances. Curr. Psychiatry Rep. 20, 59. https://doi.org/10.1007/s11920-018-0925-8.

Townsend, P., 1987. Deprivation*. J. Soc. Policy 16, 125–146. https://doi.org/10.1017/S0047279400020341.

Vancampfort, D., Hallgren, M., Firth, J., Rosenbaum, S., Schuch, F.B., Mugisha, J., Probst, M., Van Damme, T., Carvalho, A.F., Stubbs, B., 2018. Physical activity and suicidal ideation: a systematic review and meta-analysis. J. Affect. Disord. 225, 438–448. https://doi.org/10.1016/j.jad.2017.08.070.

Vancampfort, D., Stubbs, B., Mugisha, J., Firth, J., Van Damme, T., Smith, L., Koyanagi, A., 2019. Leisure-time sedentary behavior and suicide attempt among 126,392 adolescents in 43 countries. J. Affect. Disord. 250, 346–353. https://doi.org/10.1016/j.jad.2019.03.053.

Vetter, C., Chang, S.-C., Devore, E.E., Rohrer, F., Okereke, O.I., Schernhammer, E.S., 2018. Prospective study of chronotype and incident depression among middle- and older-aged women in the nurses' health study II. J. Psychiatr. Res. 103, 156. https://doi.org/10.1016/j.jpsychires.2018.05.022.

Wainberg, M., Jones, S.E., Beaupre, L.M., Hill, S.L., Felsky, D., Rivas, M.A., Lim, A.S.P., Ollila, H.M., Tripathy, S.J., 2021. Association of accelerometer-derived sleep measures with lifetime psychiatric diagnoses: a cross-sectional study of 89,205 participants from the UK biobank. PLoS Med. 18, e1003782 https://doi.org/10.1371/journal.pmed.1003782.

Walker, W.H., Walton, J.C., DeVries, A.C., Nelson, R.J., 2020. Circadian rhythm disruption and mental health. Transl. Psychiatry 10, 1–13. https://doi.org/10.1038/s41398-020-0694-0.

Wang, X., Cheng, S., Xu, H., 2019. Systematic review and meta-analysis of the relationship between sleep disorders and suicidal behaviour in patients with depression. BMC Psychiatry 19, 303. https://doi.org/10.1186/s12888-019-2302-5.

Wu, Y., Levis, B., Sun, Y., He, C., Krishnan, A., Neupane, D., Bhandari, P.M., Negeri, Z., Benedetti, A., Thombs, B.D., 2021. Accuracy of the hospital anxiety and depression scale depression subscale (HADS-D) to screen for major depression: systematic review and individual participant data meta-analysis. BMJ 373, n972. https://doi.org/10.1136/bmj.n972.

Yang, S., Berdine, G., 2017. The receiver operating characteristic (ROC) curve. SW Resp Crit Care Chron 5, 34. https://doi.org/10.12746/swrccc.v5i19.391.

Zhai, L., Zhang, H., Zhang, D., 2015. Sleep duration and depression among adults: a meta-analysis of prospective studies. Depress. Anxiety 32, 664–670. https://doi.org/10.1002/da.22386.

Zhang, M.-M., Ma, Y., Du, L.-T., Wang, K., Li, Z., Zhu, W., Sun, Y.-H., Lu, L., Bao, Y.-P., Li, S.-X., 2022. Sleep disorders and non-sleep circadian disorders predict depression: a systematic review and meta-analysis of longitudinal studies. Neurosci. Biobehav. Rev. 134, 104532 https://doi.org/10.1016/j.neubiorev.2022.104532.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B Stat Methodol. 67, 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.