Journal of Affective Disorders 363 (2024) 90–98

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research paper

The relationship between wearable-derived sleep features and relapse in Major Depressive Disorder

F. Matcham [a,b,*,1], E. Carr [c,1], N. Meyer [d], K.M. White [b], C. Oetzmann [b], D. Leightley [b], F. Lamers [e], S. Siddi [f], N. Cummins [c], P. Annas [g], G. de Girolamo [h], J.M. Haro [f], G. Lavelle [b], Q. Li [g], F. Lombardini [f], D.C. Mohr [j], V.A. Narayan [k], B.W.H.J. Penninx [d], M. Coromina [f], G. Riquelme Alacid [f], S.K. Simblett [l], R. Nica [m], T. Wykes [l,n], J.C. Brasen [g], I. Myin-Germeys [o], R.J.B. Dobson [c], A.A. Folarin [c,n], Y. Ranjan [c], Z. Rashid [c], J. Dineley [c], S. Vairavan [i], M. Hotopf [b,n], on behalf of the RADAR-CNS consortium [p]

[a] School of Psychology, University of Sussex, Falmer, UK
[b] Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[c] Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[d] Insomnia and Behavioural Sleep Medicine Clinic, University College London Hospitals NHS Foundation Trust, London, UK
[e] Department of Psychiatry and Amsterdam Public Health Research Institute, Amsterdam UMC, Vrije Universiteit, Amsterdam, the Netherlands
[f] Parc Sanitari Sant Joan de Déu, Fundació San Joan de Déu, CIBERSAM, Universitat de Barcelona, Barcelona, Spain
[g] H. Lundbeck A/S, Valby, Denmark
[h] IRCCS Instituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy
[i] Janssen Research and Development, LLC, Titusville, NJ, USA
[j] Center for Behavioral Intervention Technologies, Department of Preventative Medicine, Northwestern University, Chicago, IL, USA
[k] Davos Alzheimer's Collaborative, Wayne, PA, USA
[l] Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[m] RADAR-CNS Patient Advisory Board
[n] South London and Maudsley NHS Foundation Trust, London, UK
[o] Center for Contextual Psychiatry, Department of Neurosciences, KU Leuven, Leuven, Belgium.
[p] Radar-cns.org

ARTICLE INFO

Keywords:
Major Depressive Disorder
Sleep
Longitudinal
Wearable technology

ABSTRACT

Background: Changes in sleep and circadian function are leading candidate markers for the detection of relapse in Major Depressive Disorder (MDD). Consumer-grade wearable devices may enable remote and real-time examination of dynamic changes in sleep. Fitbit data from individuals with recurrent MDD were used to describe the longitudinal effects of sleep duration, quality, and regularity on subsequent depression relapse and severity.
Methods: Data were collected as part of a longitudinal observational mobile Health (mHealth) cohort study in people with recurrent MDD. Participants wore a Fitbit device and completed regular outcome assessments via email for a median follow-up of 541 days. We used multivariable regression models to test the effects of sleep features on depression outcomes. We considered respondents with at least one assessment of relapse ($n = 218$) or at least one assessment of depression severity ($n = 393$).
Results: Increased intra-individual variability in total sleep time, greater sleep fragmentation, lower sleep efficiency, and more variable sleep midpoints were associated with worse depression outcomes. Adjusted Population Attributable Fractions suggested that an intervention to increase sleep consistency in adults with MDD could reduce the population risk for depression relapse by up to 22 %.
Limitations: Limitations include a potentially underpowered primary outcome due to the smaller number of relapses identified than expected.

* Corresponding author at: School of Psychology, University of Sussex, Falmer, UK.
E-mail address: F.Matcham@sussex.ac.uk (F. Matcham).
[1] Joint first authors.

https://doi.org/10.1016/j.jad.2024.07.136
Received 15 December 2023; Received in revised form 9 July 2024; Accepted 16 July 2024
Available online 20 July 2024

Conclusion: Our study demonstrates a role for consumer-grade activity trackers in estimating relapse risk and depression severity in people with recurrent MDD. Variability in sleep duration and midpoint may be useful targets for stratified interventions.

1. Introduction

Major Depressive Disorder (MDD) affects 30–40 % of people at some point in their lifetime, is associated with poorer social and occupational functioning, increased physical comorbidity and premature mortality (Herrman et al., 2022), and is the leading mental health contributor to the global burden of disease (James et al., 2018). MDD follows a persistent, relapsing-remitting course in 55 % of patients and is a heterogeneous, complex condition with multiple associated bio-psychosocial factors (Verduijn et al., 2017).

Approximately 90 % of patients with MDD report sleep problems (Pandi-Perumal et al., 2020). Disrupted sleep timing and continuity and excessive sleep are core diagnostic features of depression. Insomnia, poor sleep quality, and evening chronotype are robust risk factors for the development of depression and correlate with depression severity, while morning diurnal preference has been causally associated with lower depression risk (Pandi-Perumal et al., 2020; Daghlas et al., 2021). Overlapping and bidirectional mechanisms have been implicated in both sleep-circadian function and emotional regulation (Meyer et al., 2024). Many pharmacological treatments for depression also influence sleep physiology and a range of genetic and neurotransmitter systems are implicated in MDD and sleep-circadian dysfunction (Pandi-Perumal et al., 2020). Environmental factors, including daylength and the timing of light exposure, can affect the circadian system, with implications on psychiatric risk (Burns et al., 2023).

Recent systematic reviews have shown sleep disturbances to be associated with depressive onset in both early (Scott et al., 2021) and late adulthood (Bao et al., 2017). Changes in sleep and circadian function are also leading candidate markers for early identification of relapse in MDD. Short sleep duration has been associated with an increased risk of depression recurrence (Sun et al., 2018), while greater variability in sleep-wake timing and social jet lag (discrepancies in sleep duration between work and free days) have also been associated with lower mood (Bei et al., 2017).

Mobile health (mHealth) technologies, including consumer wearables and smartphones, have increasingly attracted interest as tools for remote and real-time examination of dynamic changes in mood, activity, and sleep variables (Matcham et al., 2019; Zhang et al., 2021). Consumer sleep technologies are widely available, affordable, and acceptable and offer longitudinal sleep monitoring in an individual's typical sleep environment. They circumvent many of the limitations of conventional objective sleep and rest-activity measures: polysomnography is costly, cross-sectional, and typically only available in laboratory settings, while most clinical-grade actigraphy devices require data to be downloaded from the device manually, restricting long-term use.

Consumer wearables, therefore, represent a promising alternative to traditional actigraphy (de Zambotti et al., 2023). Studies comparing consumer sleep technologies with clinical/research-grade actigraphy have demonstrated equivalent or better performance of consumer technologies in healthy populations (Lee et al., 2019; Chinoy et al., 2021) and those with insomnia (Kahawage et al., 2020; Hamill et al., 2020). Compared to polysomnography, consumer devices perform similarly to clinical/research-grade actigraphy, with high sensitivity and moderate specificity for sleep/wake classification (de Zambotti et al., 2018; Tedesco et al., 2019; Chinoy et al., 2021). mHealth holds considerable potential for developing Just-In-Time Adaptive Interventions, using dynamic, individualised data to personalise the timing and content of sleep interventions.

Remote Assessment of Disease and Relapse – Major Depressive Disorder (RADAR-MDD; Matcham et al., 2019) was an international, multicentre cohort study in people with recurrent MDD that leveraged data from smartphones and wearable devices. The design of RADAR-MDD was driven by extensive research with users with lived experience of MDD and Patient and Public Involvement to maximise participant uptake and long-term engagement with the project (Simblett et al., 2019; Polhemus et al., 2020). RADAR-MDD identified the appearance, form factor, and cost of clinical-grade actigraphy devices as significant limitations for longitudinal studies with multi-year follow-up. The study is the largest remote measurement study in depression conducted to date.

We have previously reported short-term associations of sleep features with depressive symptomatology measured via app-delivered questionnaires (Zhang et al., 2021). Here, we extended our previous work in several ways. First, by leveraging the longitudinal design of RADAR-MDD to assess within-individual effects of sleep on depression, accounting for between-individual differences. Second, by considering longer-term effects over several months rather than several weeks. Third, by considering clinically-assessed depressive relapse in addition to depression severity. Our analysis considered the effects of (1) sleep duration, (2) sleep quality, and (3) sleep regularity on depression relapse and severity.

2. Methods

2.1. Study design and participants

This paper presents a secondary analysis of RADAR-MDD (Matcham et al., 2019), a dataset following people with recurrent MDD for a median of 541 days (interquartile range (IQR): 401–730 days; Matcham et al., 2022). Participants were recruited from the UK, Netherlands and Spain, with ethical approvals from the Camberwell St Giles Research Ethics Committee in London (reference: 17/LO/1154), CEIC Fundacio Sant Joan de Deu in Barcelona (CI: PIC-128-17) and the Medische Ethische Toetsingscommissie VUmc in Amsterdam (METcVUmc registratienummer 2018.012 – NL63557.029.17).

Eligible participants had at least two MDD episodes with at least one episode in the previous two years; were able to complete self-reported questionnaires via a smartphone; were fluent in English, Dutch, Catalan or Spanish; were willing and able to give informed consent; were an Android user or willing to switch to an Android phone for the study; and were aged >18 years. Individuals were excluded if they had a history of bipolar disorder, schizoaffective disorder, schizophrenia, MDD with psychotic features, dementia, recent drug or alcohol misuse or a major medical illness (requiring long periods of hospitalisation). Recruitment occurred between November 2017 and June 2020, and follow-up ceased on 30th April 2021.

2.2. Patient and public involvement

RADAR-MDD was co-developed with service users in our Patient Advisory Board (PAB). They were involved in the choice of measures, the timing and issues of engagement, and the development of the analysis plan. Representatives of the PAB were invited to author and critically review this paper.

2.3. Procedure

Participants were asked to wear a wrist-worn Fitbit Charge device and complete regular questionnaires throughout follow-up, administered via email using the Research Electronic Data Capture (REDCap)

platform (Harris et al., 2009). Participants enrolled before September 2019 received a Charge 2 device; all participants enrolled subsequently were offered a Charge 3 device since the previous model was discontinued by Fitbit. Questionnaires assessed sociodemographic factors, medical and psychiatric history, service and medication use, health behaviours and clinical characteristics questionnaires (Supplementary Material 1).

2.4. Depression outcomes

We considered two outcomes assessed repeatedly every three months during follow-up:

1) **MDD relapse** was defined as meeting three criteria:

   i. Meeting the World Health Organisation Composite International Diagnostic Interview – Short Form criteria for MDD (CIDI-SF). The CIDI-SF MDD criteria required individuals to endorse at least five of the nine MDD symptoms, one of which must be a core symptom of low mood or anhedonia. The CIDI-SF has excellent sensitivity and specificity in identifying current MDD state and is used extensively on web-based platforms (Kessler et al., 1998).
   ii. Scoring >25 on the IDS-SR (Inventory for Depressive Symptomatology – Self Report; Trivedi et al., 2004) indicating at least moderate symptom severity.
   iii. Having been in a state of remission within the past six months.

A trained research worker called participants who met the above criteria to confirm their relapse status and questionnaire responses, assess the timing of relapse onset, and conduct risk assessments where appropriate. For this outcome only, we excluded participants who did not meet the CIDI-SF MDD criteria but scored over 25 on the IDS-SR (indicating chronically severe depression symptoms) to ensure the comparison group ('Non-relapse') included only those with low depression severity.

2) **Depressive symptom severity** was assessed based on total scores on the IDS-SR. Scores range from 0 to 80, and higher scores indicate increased depression severity.

2.5. Sleep features

The Fitbit device tracked sleep using accelerometer and photoplethysmography sensors. Participants were asked to wear the device continuously, removing it only to charge the battery and when showering or swimming. The device has comparable performance to research-grade actigraphy devices in estimating sleep duration, sleep efficiency and sleep onset latency (de Zambotti et al., 2018, 2023; Tedesco et al., 2019).

Fitbit data were gathered continuously and summarised as daily indicators (e.g., 'Total hours of sleep'). For each day, we focused on the primary sleep event, defined as the longest continuous sleep period, thereby excluding shorter sleep events such as napping. We considered 13 sleep features split into three domains: (1) duration, (2) quality, and (3) regularity (Table 1). For each feature, we aggregated the daily measures for the four weeks before each 3-monthly outcome assessment. We also considered changes in sleep between successive outcome assessments (see Methodological Supplement). We used a four-week reference period for sleep following discussions with sleep consultants and psychiatrists, as well as past studies showing deteriorations in sleep quality in the four weeks before relapse (Young et al., 1991; Fang et al., 2019). We were also motivated by the timing of potential future interventions. Four weeks was considered the minimum period needed for future screening tools to detect sleep disruptions and initiate an intervention. Within each four-week period, we included participants providing sleep information on at least 8 days. We chose a minimum of 8 days to derive reliable summaries while minimising selection bias (participants wearing their Fitbit for more than eight days tended to have lower depressive symptoms and more stable sleep patterns; Sun et al., 2023).

Table 1
Sleep features measured in the four weeks prior to outcome assessment, derived from daily summaries of continuously-collected Fitbit indicators. Continuous Fitbit.

Domain	Feature	Definition
Sleep duration	Total sleep time[a]	Median daily hours spent asleep.
	Change in total sleep time[b]	Difference in median daily hours spent asleep over the four weeks before the current outcome assessment (T_2) versus the four weeks before to the previous outcome assessment three months ago (T_1).
	Total sleep time, variance[a]	Intra-individual variance in median daily hours spent asleep.
	Change in total sleep time, variance[b]	Difference in intra-individual variance in median daily hours asleep from T_2 to T_1.
Sleep quality	Sleep efficiency[a]	Median hours spent asleep as a proportion of hours spent in bed per day.
	Sleep fragmentation index[a]	Median daily number of awakenings divided by the hours spent asleep.
	Sleep onset latency[a]	Median time between last recorded step taken and Fitbit-derived sleep onset each day.
	Sleep onset latency variance[a]	Intra-individual variance in median sleep onset.
Sleep regularity	Sleep midpoint[a]	Median midpoint between sleep onset and sleep offset.
	Sleep midpoint, variance[a]	Intra-individual variance in median sleep midpoint.
	Change in sleep midpoint[b]	Difference in median sleep midpoint from T_2 to T_1.
	Change in sleep midpoint variance[b]	Difference in intra-individual variance in median sleep midpoint from T_2 to T_1.
	Social jet lag[a]	Absolute difference between weekday and weekend sleep midpoints.

[a] Measured over the four weeks prior to each outcome assessment.
[b] Measured as the change between the four weeks before the current outcome assessment (T_2) and the four weeks before the previous outcome assessment three months ago (T_1).

2.6. Covariates

We considered covariates previously shown to affect sleep and depression (Bei et al., 2016), including variables measured at enrolment (age, gender, years of education and partnership status) and at each 3-monthly outcome assessment (atypical depression subtype, medication use, alcohol use, and hours of daylight in the previous month, a proxy for seasonal effects). Atypical depression is characterised by hypersomnia and increased appetite and has markedly different sleep phenotypes compared to other subtypes (Posternak, 2003). Atypical depression subtype was defined as mood reactivity between 0 and 2 and at least two symptoms from leaden paralysis, weight gain, increased appetite, hypersomnia, and interpersonal sensitivity (Novick et al., 2005). Medication use was measured using three binary variables capturing medications related to depression (antidepressants, antipsychotics, anticonvulsant, stimulating antidepressant), sleep (benzodiazepine, hypnotic) or 'other'. Alcohol use was measured using the Alcohol Use Disorders Identification Test (AUDIT; Daeppen et al., 2000) total score (see Methodological Supplement).

2.7. Statistical analyses

The analyses were conducted in three parts. First, we described outcomes and covariates using appropriate summary statistics. Second, we tested the effect of each sleep feature on each depression outcome by

fitting a series of binary logistic (relapse) and linear (severity) multivariable regression models. Each sleep feature was tested in a separate model. All models included (i) a participant-level random intercept to account for the clustering of repeated outcome assessments, (ii) the covariates listed above, (iii) quadratic terms for the sleep feature to allow for non-linear effects on depression, (iv) the participant's mean value for the sleep feature to estimate within-person effects (see below). Models for depression severity additionally included the previous value of the outcome (at the previous three-monthly assessment) as both a linear and quadratic term, recognising that prior severity may have non-linear effects on current severity (Vittengl et al., 2016).

We summarised the effect of each sleep feature on depression using average marginal effects and adjusted predictions. The marginal effect is the partial derivative of the regression equation with respect to each variable in the model for each unit in the data. The average marginal effect (AME) is the mean of these partial derivatives over the sample. For relapse, the AME represents the percentage points change in the probability of relapse per standard deviation difference in the sleep feature; for depression severity, the change in the IDS-SR total score (compared to three months ago) per standard deviation difference in the sleep feature. For selected sleep features, we additionally plotted adjusted predictions. These represent the model-predicted value of the outcome for a range of values for each sleep feature (from −2 SD to +2 SD), holding other variables to their median values.

These models were estimated in a Bayesian framework using Stan and the brms package for R (Bürkner, 2017). All models were sampled for 40,000 iterations and thinned by retaining every 10th sample. We used weakly informative priors, centred on zero, to constrain estimates to plausible values and serve as a form of statistical regularisation by shrinking coefficients towards zero (Lemoine, 2019; see Methodological Supplement). No corrections were made for multiple testing (Sjölander and Vansteelandt, 2019). Posterior draws for AMEs were summarised as the median and 50 % and 89 % credible intervals. We retained all available outcome assessments under the missing at random assumption.

Third, to illustrate the potential for future sleep interventions, we calculated adjusted population attributable fractions (PAF) for depression relapse. The PAF represents the fraction of cases in the population that would be prevented if a specific exposure was eliminated (Mansournia and Altman, 2018). Since sleep features were measured continuously, we dichotomised each feature to create a binary exposure where '1' represents the upper quartile and '0' represents the lower three quartiles. The PAF, therefore, represents the fraction of relapse cases prevented by an intervention that eliminates a dichotomised sleep exposure such as 'high sleep variability' or 'low sleep efficiency'. We calculated adjusted PAFs from binary logistic regression models using the glm and AFglm functions in R.

2.7.1. Centring and standardisation of sleep features

To estimate the within-person effects of sleep on depression, we (i) participant-mean-centred each sleep feature by subtracting the mean of the participant's repeated assessments and (ii) included in each model a term representing the participant's mean value (Curran and Bauer, 2011). Sleep features were additionally standardised (standard deviation = 0), and some were log-transformed (see Methodological Supplement for details).

2.8. Sensitivity analyses

We conducted two sensitivity analyses. First, we investigated whether the effect of sleep differed for participants with atypical depression by comparing models with and without an interaction term (sleep feature × atypical depression subtype). Models were compared using 10-fold cross-validation (Vehtari et al., 2016) implemented in the loo package for R (Vehtari et al., n.d.). Interaction terms were considered significant if the expected log pointwise predictive density utility score (ELPD) was ≥4 and at least twice the ELPD standard error (Sivula et al., 2022). Second, we repeated the models for depression severity after removing four sleep items from the IDS-SR to ensure that any associations were not the result of overlapping measures.

3. Results

Of 623 participants in the RADAR-MDD sample, we excluded 227 participants who did not provide any information on sleep, two participants without any outcome assessments, and one participant with missing covariate information. Excluded participants were similar to those analysed in terms of age, gender, partnership status, and education (Supplementary Table 2.1). We have previously published recruitment rates, retention rates, and data availability for the RADAR-MDD study (Matcham et al., 2022).

We derived two analytical samples comprising people with information on depression relapse (n = 218 individuals; 624 assessments) and depression severity (n = 393 individuals; 1361 assessments), respectively. The analytical samples overlapped: all participants (n = 393) had information on depression severity, and a subset of these (n = 218) also had information on relapse. Table 2 presents participant characteristics at enrolment. Participants had a median age of 49–52 years, and 24 % were male. Around half were living with a partner, and two-thirds were taking an antidepressant medication. For depression severity, the median (IQR) IDS-SR score was 30 (20, 42); relapse was recorded at 39/624 (6.2 %). Supplementary Table 3 presents characteristics of the depression severity sample (n = 393) stratified by completion rates, showing that participants who completed more outcome assessments tended to be older, female, and less likely to be taking an antidepressant at enrolment. Fig. 1 presents AMEs for a one standard deviation (SD) difference in each sleep feature (see Supplementary Table 2.2).

Table 2
Characteristics of the analytical samples.

	Depression relapse[a]	Depression severity[a]
	n = 218 individuals	n = 393 individuals
Participant characteristics at enrolment		
Age[b]	52.0 (34.0, 58.9)	49.0 (32.0, 57.0)
Years of education[b]	17.0 (14.0, 19.0)	16.0 (13.0, 18.0)
AUDIT total score[b]	3.0 (1.0, 5.0)	2.0 (0.0, 4.0)
Male gender[c]	52 (24)	93 (24)
Lives with partner[c]	125 (57)	198 (50)
Medication (depression)[c]	137 (63)	269 (68)
Medication (sleep)[c]	21 (10)	72 (18)
Medication (other)[c]	5 (2)	8 (2)
Outcomes over follow-up		
No. follow-up assessments	624	1361
Relapse[d]	39 (6.2)	–
IDS-SR total score[b]	17.0 (12.0, 21.0)	28.0 (18.0, 37.0)

[a] There were two overlapping analytical samples. The sample for depression severity (n = 393) comprises participants with at least one assessment of the IDS-SR score. The sample for depression relapse is a subset of participants (n = 218) who additionally had at least one assessment of relapse. Depression severity scores in the relapse sample are lower due to the definition of the comparison group (i.e., 'Non-relapse'), which excluded participants who had not relapsed but still had high severity (IDS-SR over 25).
[b] Median (IQR).
[c] N (%).
[d] Information on relapse was unavailable for some participants in the severity sample.

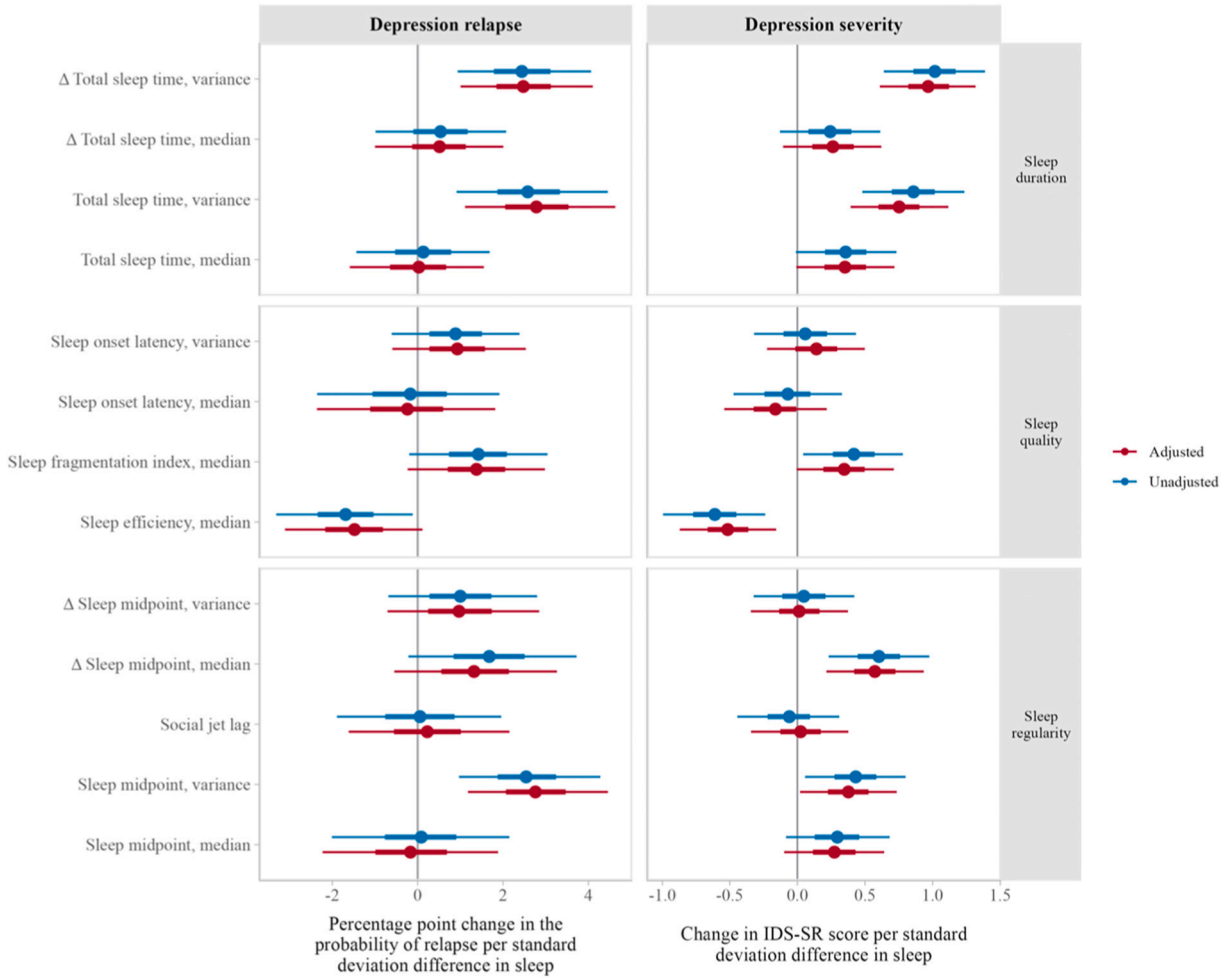

Fig. 1. Average marginal effects for a 1 SD difference in each sleep feature.
Notes. This figure presents the posterior distributions of the average marginal effects (AMEs) for each sleep feature and outcome. The points represent the median of the posterior distribution. The estimates represent the within-person effects, derived by person-mean centring each sleep feature and including the participant's mean value (across repeated follow-up assessments) as a covariate in the model. The thick and thin lines represent the 50 % and 89 % credible intervals, respectively. The estimates are presented unadjusted (in red) and adjusted (in blue) for covariates measured at enrolment (age, gender, years of education, partnership status) and during follow-up (atypical depression subtype, medication use, alcohol use, hours of daylight). The AMEs are presented unadjusted (in red) and adjusted (in blue). "Δ" refers to the change in the respective sleep feature between two consecutive outcome assessments, three months apart (see Methodological Supplement for details).
For depression relapse, the AMEs represent the percentage point change in the probability of relapse per 1 SD difference in the respective sleep feature. For example, for 'total sleep time variance', the AME was 2.8, indicating that the probability of relapse was 2.8 % higher at assessments where participants' sleep time variability was one standard deviation higher compared to their variability at other assessments. Since relapse was observed at just 6.2 % of outcome assessments, an increase of 2.8 % represents almost a 50 % relative increase in the probability of relapse. For depression severity, the AME represents the units change in IDS-SR severity since the last 3-monthly outcome assessment per standard deviation difference in the respective sleep feature. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.1.1. Sleep duration

Greater intra-individual variability in total sleep time was associated with an increased probability of relapse and increased depression severity. A one standard deviation increase in 'total sleep time variance' was associated with a 2.8 % increase in the probability of relapse (89 % Credible Interval (CrI): 1.1, 4.6) and a 0.8 (89 % CrI: 0.4, 1.1) unit increase in depression severity. Greater sleep duration was associated with small increases in depression severity (0.4 unit increase in severity per 1 SD difference in sleep duration; CrI: 0.0, 0.7), but we found no evidence of an effect on depression relapse.

3.1.2. Sleep quality

Higher sleep fragmentation was associated with an increased probability of relapse (AME = 1.4; CrI: −0.2, 3.0) and higher depression severity (0.3; 89 % CrI: −0.1, 0.7). Conversely, higher sleep efficiency was associated with a reduced probability of relapse (−1.5; CrI: −3.1, 0.1) and lower depression severity (−0.5; CrI: −0.9, −0.2). We found no evidence of an effect of sleep onset latency on either outcome.

3.1.3. Sleep regularity

More variable sleep midpoints were associated with an increased probability of relapse (AME = 2.8; CrI: 1.2, 4.5) and higher depression severity (AME = 0.4; CrI: 0.0, 0.7). Increases in sleep midpoints between successive outcome assessments were similarly associated with an increased probability of relapse and higher severity, but the credible intervals for relapse crossed zero, suggesting no effect.

Fig. 2 presents adjusted predictions for selected sleep features. These represent model-based predictions generated across a range of values for each sleep feature while holding other covariates to their median values.

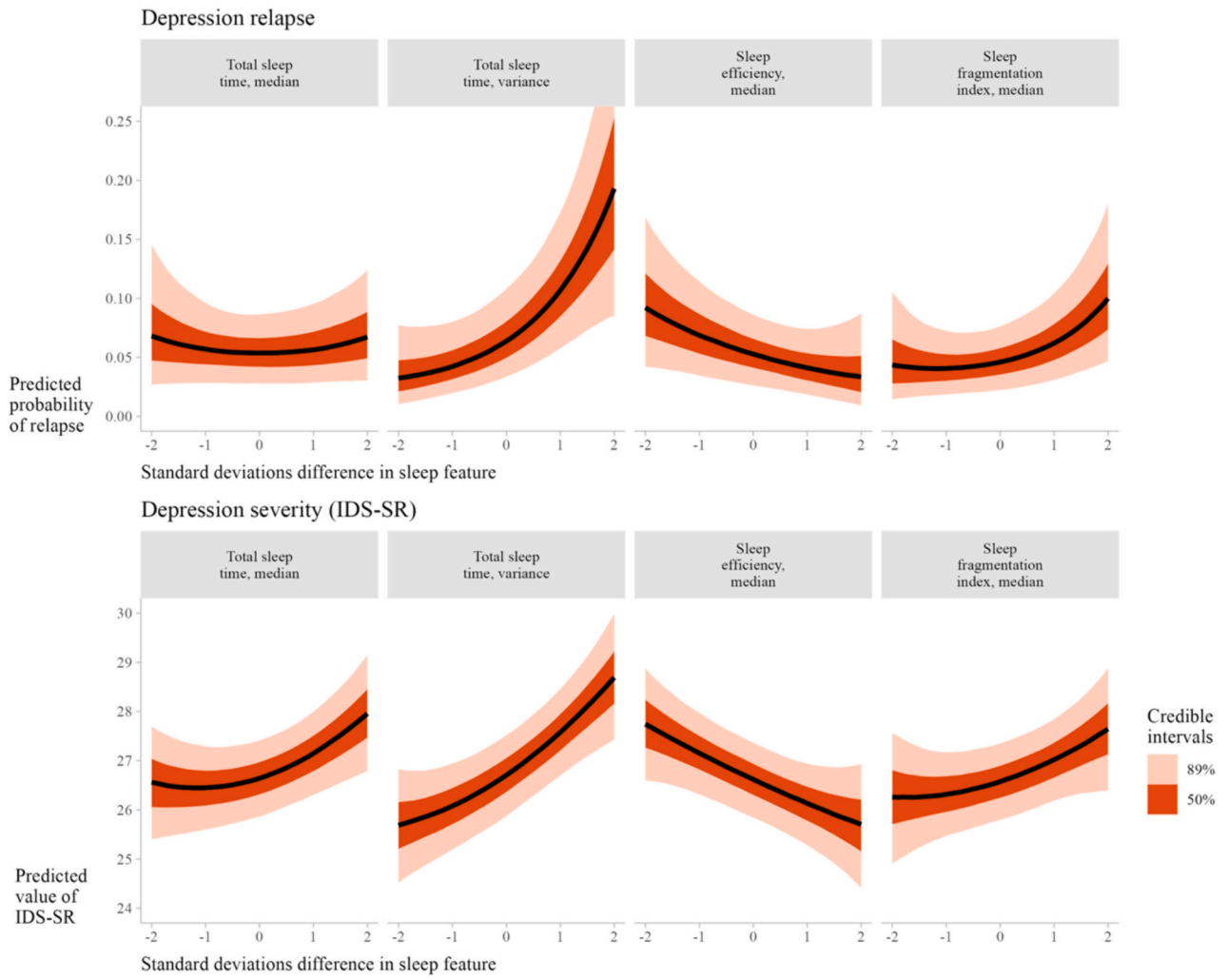

Fig. 2. Adjusted predictions for selected sleep features.
Notes. This figure presents model-based predictions for each outcome, for a range of values for each sleep feature (−2 to +2 standard deviations), while holding other covariates to their median values. The black line represents the median of the posterior distribution; the darker and lighter regions represent the 50 % and 89 % credible intervals, respectively.

For both outcomes, we found that participants with more variable sleep tended to experience poorer outcomes. For example, the predicted probability of relapse for participants with average sleep time variance was ~6 %, compared to 10–15 % for those with variances 1–2 SD above average. Conversely, participants with greater sleep efficiency were less likely to relapse and tended to report lower depression severity. For example, from a predicted probability of relapse of ~6 % among those with average sleep efficiency to 3–4 % among those with efficiency 1–2 SD above average.

3.1.4. Adjusted population attributable fractions

Presented in Supplementary Table 2.3, most adjusted PAFs were small and had confidence intervals that crossed zero. For the dichotomised measure of 'change in sleep time variance' (1 = upper quartile of participants; 0 = lower quartiles), the PAF was 22 % (95 % Confidence Interval (CI): 3 %, 41 %). This suggests that if an intervention could eliminate this exposure (i.e., to eliminate large increases in sleep time variability), then 22 % of depression relapse cases in the population could be prevented.

3.1.5. Sensitivity analyses

We found no evidence of differences in the effect of sleep for participants with atypical depression. For all models, adding an interaction term (sleep feature × atypical subtype) resulted in negligible improvements in model fit (Supplementary Table 2.4). Our findings were also consistent when removing sleep items from IDS-SR, although some effect sizes were slightly attenuated (Supplementary Table 2.2).

4. Discussion

We found that increased variability in sleep duration and sleep timing regularity, and greater sleep fragmentation were associated with elevated depression severity and relapse risk. Conversely, we found negative associations between sleep efficiency and depression outcomes. We have previously reported short-term associations between sleep features and depressive symptoms in RADAR-MDD (Zhang et al., 2021). The present analysis expands on earlier findings by estimating within-participant effects, including effect size estimates, more robust depression outcomes, considering non-linear effects, and accounting for important confounders. Our findings suggest that moderate increases in variability in sleep parameters are associated with clinically meaningful increases in the probability of relapse. For example, greater sleep variability was associated with increases in the probability of relapse from 6 % to 9 %, a 50 % relative increase. While subject to several limitations listed below, these effect sizes nonetheless highlight the potential of interventions targeting sleep in MDD cohorts.

The Lancet Psychiatry *Commission on Psychological Treatments Research in Tomorrow's Science* highlights the potential for technology to deliver novel interventions (Holmes et al., 2018). Our findings demonstrate the potential of consumer technologies to measure sleep parameters that may be predictive of deterioration and relapse in MDD and thus provide targets for personalised intervention. These findings broaden the existing evidence base (Pandi-Perumal et al., 2020) to include measurements of *variability* in sleep parameters. Accurate assessments of sleep variability are challenging to capture using the conventional use of short-term PSG or Experience Sampling Method (ESM), which often exhibits subjective reporting bias (Palmier-Claus et al., 2011). Our findings also highlight the importance of within-participant variability for predicting depression outcomes up to three months into the future, over and above absolute sleep duration. Furthermore, due to the longitudinal nature of the dataset, we were able to show that more variable sleep patterns preceded relapse and may, therefore, be targetable in interventions to improve depression outcomes.

Our work has limitations. First, when planning RADAR-MDD, we anticipated observing 100 relapses during follow-up (Trivedi et al., 2006), whereas our analysis included only 39 relapse events. One explanation for this is our inclusive eligibility criteria: RADAR-MDD aimed to identify predictors of relapse *and* remission and, therefore, did not exclude participants based on their depression severity at enrolment. This meant many participants were unable to experience a relapse due to their pre-existing high depression severity at enrolment. The consistency of our findings for the two depression outcomes, with effects in similar directions for all sleep features, suggests that the absence of evidence for some relapse models may represent a lack of statistical power rather than a lack of association.

Second, we required participants to contribute sleep information for at least 8 days in the four weeks preceding each outcome assessment. This minimum was chosen to balance the need for reliable summaries of sleep parameters against the number of excluded observations. Third, firmware used by Fitbit devices varied over follow-up. When the Charge 2 device was discontinued 18 months into recruitment, we made the pragmatic decision to start using the Charge 3 device instead, and there were numerous firmware updates during the follow-up period. The changes in proprietary algorithms used to determine heart rate may have influenced our results (Nelson et al., 2020). Fourth, to calculate PAFs, we dichotomised each sleep feature to identify participants in the upper quartile. This was a pragmatic decision to illustrate the potential of future interventions but is a poor representation of the continuous sleep features. Finally, the follow-up period for RADAR-MDD (2017 to 2021) partially overlapped with the COVID-19 pandemic (from March 2020). While previous analyses of this cohort suggest patterns of depressive symptoms and sleep were largely stable during this period (Leightley et al., 2021), the available sample size meant we were unable to examine differences from before and after the pandemic onset.

There is an ongoing debate about the accuracy of optical heart rate measurements using photoplethysmography across different skin tones (de Zambotti et al., 2023). Darker skin tones absorb more green light and, therefore, reduce the accuracy of heart rate estimation, which is used in sleep/wake classification (Bent et al., 2020). However, evidence is conflicting, with a recent validation study comparing a range of devices (including the Fitbit Charge 2) reporting no statistically significant differences in accuracy across different skin tones (Colvonen, 2021). These findings relied on a small sample and an imperfect skin tone measurement, however. A recent systematic review highlighted the need for higher-quality evidence on the accuracy of wearable devices among minority ethnic groups and the subsequent impacts on health and care disparities (Koerber et al., 2022). Due to differing ethical committee requirements, we could not collect ethnicity data from two of the three countries we recruited from. This prevented us from investigating the role of ethnicity or from including it as a covariate in our analyses and limits the generalizability of our findings for individuals with different skin tones.

Furthermore, the high physical and mental health comorbidity of our study population reflects the frequent association between depression and other illnesses in clinical and epidemiological studies (Von Korff et al., 2009). The differences in the exact patterns of comorbidity between our sample and clinical populations may affect the generalizability of our findings. The high rate of comorbidity may further have introduced a degree of "noise" in our models for the association between sleep parameters and depressive relapse, potentially weakening the associations we report.

Our analyses required participants to use a wrist-worn wearable device and an Android smartphone for up to two years. This introduced potential selection effects, as some individuals may have declined participation to avoid switching phones or because they weren't comfortable with remote data collection. Further work is needed to understand barriers to uptake for mHeath studies and interventions (Oetzmann et al., 2022).

A strength of the study is the inclusion of covariates, including atypical depression, which has a unique sleep phenotype, and seasonal changes in daylength. While we made every effort to adjust our models appropriately, aspects such as medication use are complex and challenging to capture effectively longitudinally. We relied on self-reported medication use at each time point and coded medications considered to impact sleep or depression, an imperfect method which limits information about changes in medication over time.

Future analyses should extend this work by developing multivariable prediction models for depression relapse based on sleep and other wearable biomarkers. Such models could explore alternative windows for sleep aggregation (e.g., 1, 2, or 3 weeks) besides the 4-week window used in our analyses. Robust, validated predictive models would enable stratified interventions targeting individuals with more variable sleep duration or timing. Our analyses suggested that effective interventions to reduce sleep variability could reduce around 20 % of the population risk for depression relapse. While this estimate relies on strong causal assumptions and presumes the availability of an intervention that can eliminate the exposure, it nonetheless illustrates the potential of future interventions. Our analysis focused on the sleep-depression relationship, mirroring the anticipated clinical implementation where interventions would target sleep (duration, quality, or regularity) with the aim of reducing subsequent relapse events. Nevertheless, it is crucial for future studies to explore the reciprocal association, acknowledging the probable existence of bi-directional relationships between sleep and depression.

Disordered sleep is a prevalent and disruptive feature of depression that is challenging to capture. Our study demonstrates a role for commercially available activity trackers in estimating relapse risk and depression severity in people with established, recurrent MDD. Specifically, variability in sleep duration and midpoint may be useful targets for stratified interventions.

CRediT authorship contribution statement

F. Matcham: Writing – review & editing, Writing – original draft, Visualization, Supervision, Project administration, Methodology, Investigation, Conceptualization. **E. Carr:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **N. Meyer:** Writing – review & editing, Methodology, Conceptualization. **K.M. White:** Writing – review & editing, Investigation. **C. Oetzmann:** Writing – review & editing, Investigation. **D. Leightley:** Writing – review & editing, Validation, Resources, Data curation. **F. Lamers:** Writing – review & editing, Supervision, Resources, Project administration, Investigation. **S. Siddi:** Writing – review & editing, Resources, Project administration, Investigation. **N. Cummins:** Writing – review & editing, Validation, Data curation. **P. Annas:** Supervision, Project administration, Investigation. **G. de Girolamo:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition. **J.M. Haro:**

Writing – review & editing, Resources, Project administration, Funding acquisition. **G. Lavelle:** Writing – review & editing, Supervision, Investigation. **Q. Li:** Writing – review & editing, Investigation. **F. Lombardini:** Writing – review & editing, Investigation. **D.C. Mohr:** Writing – review & editing, Investigation, Funding acquisition. **V.A. Narayan:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **B.W.H.J. Penninx:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition. **M. Coromina:** Writing – review & editing, Investigation. **G. Riquelme Alacid:** Writing – review & editing, Investigation. **S.K. Simblett:** Writing – review & editing, Conceptualization. **R. Nica:** Writing – review & editing, Conceptualization. **T. Wykes:** Writing – review & editing, Funding acquisition, Conceptualization. **J.C. Brasen:** Writing – review & editing, Conceptualization. **I. Myin-Germeys:** Writing – review & editing, Funding acquisition. **R.J.B. Dobson:** Writing – review & editing, Software, Resources, Funding acquisition. **A. A. Folarin:** Writing – review & editing, Software, Funding acquisition. **Y. Ranjan:** Writing – review & editing, Software. **Z. Rashid:** Writing – review & editing, Software, Conceptualization. **J. Dineley:** Writing – review & editing, Validation, Data curation. **S. Vairavan:** Writing – review & editing, Data curation. **M. Hotopf:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

HE was a full-time employee of H. Lundbeck A/S and held stock and stock options in H. Lundbeck A/S at the time of study conduct. QL, SV and VN were employees of Janssen Research & Development, LLC and held company stocks/stock options at the time of study conduct. JMH has received economic compensation for participating in advisory boards or giving educational lectures from Eli Lilly & Co, Sanofi, Lundbeck, and Otsuka. No other authors have competing interests to declare.

Acknowledgements

The RADAR-CNS project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115902. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA (www.imi.europa.eu). This paper represents independent research part-funded by the NIHR Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Participant recruitment in Amsterdam was partially accomplished through Hersenonderzoek.nl, a Dutch online registry that facilitates participant recruitment for neuroscience studies (https://hersenonderzoek.nl/). Hersenonderzoek.nl is funded by ZonMw-Memorabel (project no 73305095003), a project in the context of the Dutch Deltaplan Dementie, Gieskes-Strijbis Foundation, the Alzheimer's Society in the-Netherlands and Brain Foundation Netherlands. Participants in Spain were recruited through the following in-stitutions: Parc Sanitari Sant Joan de D´eu network of mental health services (Barcelona); Institut Catal'a de la Salut primary care services (Barcelona); Institut Pere Mata-Mental Health Care (Terrassa); Hospital Clínico San Carlos (Madrid). We thank all the members of the RADAR-CNS patient advisory board for their contribution to the device selection procedures, and their invaluable advice throughout the study protocol design. This research was reviewed by a team with experience of mental health problems and their carers who have been specially trained to advise on research proposals and documentation through the Feasibility and Acceptability Support Team for Researchers (FAST-R): a free, confidential service in England provided by the National Institute for Health Research Maudsley Biomedical Research Centre via King's College London and South London and Maudsley NHS Foundation Trust. We thank all GLAD Study volunteers for their participation, and gratefully acknowledge the NIHR BioResource, NIHR BioResource centres, NHS Trusts and staff for their contribution. We also acknowledge NIHR BRC, King's College London, South London and Maudsley NHS Trust and King's Health Partners. We thank the National Institute for Health Research, NHS Blood and Transplant, and Health Data Research UK as part of the Digital Innovation Hub Programme.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jad.2024.07.136.

References

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.

Bei, B., Wiley, J.F., Trinder, J., Manber, R., 2016. Beyond the mean: a systematic review on the correlates of daily intraindividual variability of sleep/wake patterns. Sleep Med. Rev. 28, 108–124. https://doi.org/10.1016/j.smrv.2015.06.003.

Bei, B., Manber, R., Allen, N.B., Trinder, J., Wiley, J.F., 2017. Too long, too short, or too variable? Sleep intraindividual variability and its associations with perceived sleep quality and mood in adolescents during naturalistically unconstrained sleep. Sleep 40 (2). https://doi.org/10.1093/sleep/zsw067.

Bent, B., Goldstein, B.A., Kibbe, W.A., Dunn, J.P., 2020. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Dig. Med. 3 (18) https://doi.org/10.1038/s41746-020-0226-6.

Bürkner, P.-C., 2017. Brms: an R package for Bayesian multilevel models using Stan. J. Stat. Softw. 80 (1), 1–28. https://doi.org/10.18637/jss.v080.i01.

Burns, A.C., Windred, D.P., Rutter, M.K., Olivier, P., Vetter, C., Saxena, R., Lane, J.M., Phillips, A.J.K., Cain, S.W., 2023. Day and night light exposure are associated with psychiatric disorders: an objective light study in >85,000 people. Nat. Ment. Health 1, 853–862. https://doi.org/10.1038/s44220-023-00135-8.

Chinoy, E.D., Cuellar, J.A., Huwa, K.E., Jameson, J.T., Watson, C.H., Bessman, S.C., Hirsch, D.A., Cooper, A.D., Drummond, S.P.A., Markwald, R.R., 2021. Performance of seven consumer sleep-tracking devices compared with polysomnography. Sleep 44, zsaa291. https://doi.org/10.1093/sleep/zsaa291.

Colvonen, P.J., 2021. Response to: investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Dig. Med. 4 (38) https://doi.org/10.1038/s41746-021-00408-5.

Curran, P.J., Bauer, D.J., 2011. The disaggregation of within-person and between-person effects in longitudinal models of change. Annu. Rev. Psychol. 62, 583–619. https://doi.org/10.1146/annurev.psych.093008.100356.

Daeppen, J.B., Yersin, B., Landry, U., Pecoud, A., Decrey, H., 2000. Reliability and validity of the alcohol use disorders identification test (AUDIT) imbedded within a general health risk screening questionnaire: results of a survey in 332 primary care patients. Alcohol. Clin. Exp. Res. 24 (5), 659–665. https://doi.org/10.1111/j.1530-0277.2000.tb02037.x.

Daghlas, I., Lane, J.M., Saxena, R., Vetter, C., 2021. Genetically proxied diurnal preference, sleep timing, and risk of Major Depressive Disorder. JAMA Psychiatry 78 (8), 903–910. https://doi.org/10.1001/jamapsychiatry.2021.0959.

de Zambotti, M., Goldstone, A., Claudatos, S., Colrain, I.M., Baker, F.C., 2018. A validation study of Fitbit charge 2™ compared with polysomnography in adults. Chronobiol. Int. 35 (4), 465–476. https://doi.org/10.1080/07420528.2017.1413578.

de Zambotti, M., Goldstein, C., Cook, J., Menghini, L., Altini, M., Cheng, P., Robillard, R., 2023. State of the science and recommendations for using wearable technology in sleep and circadian research. Sleep, zsad325. https://doi.org/10.1093/sleep/zsad325.

Fang, H., Tu, S., Sheng, J., Shao, A., 2019. Depression in sleep disturbance: a review on a bidirectional relationship, mechanisms and treatment. J. Cell. Mol. Med. 23 (4), 2324–2332. https://doi.org/10.1111/jcmm.14170.

Hamill, K., Jumabhoy, R., Kahawage, P., de Zambotti, M., Walters, E.M., Drummond, S.P.A., 2020. Validity, potential clinical utility and comparison of a consumer activity tracker and a research-grade activity tracker in insomnia disorder II: outside the laboratory. J. Sleep Res. 29, e12944 https://doi.org/10.1111/jsr.12944.

Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., Conde, J.G., et al., 2009. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J. Biomed. Inform. 42 (2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010.

Herrman, H., Patel, V., Kieling, C., Berk, M., Buchweitz, C., Cuijpers, P., et al., 2022. Time for united action on depression: a Lancet–World Psychiatric Association Commission. Lancet 399 (10328), 957–1022. https://doi.org/10.1016/S0140-6736(21)02141-3.

Holmes, E.A., Ghaderi, A., Harmer, C.J., Ramchandani, P.G., Cuijpers, P., Morrison, A.P., et al., 2018. Commission on psychological treatments research in tomorrow's

science. Lancet Psychiatr. 5 (3), 237–286. https://doi.org/10.1016/S2215-0366(17)30513-8.
James, S.L., Abate, D., Abate, K.H., Abay, S.M., Abbafati, C., Abbasi, N., et al., 2018. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. Lancet 392 (10159), 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7.
Kahawage, P., Jumabhoy, R., Hamill, K., de Zambotti, M., Drummond, S.P.A., 2020. Validity, potential clinical utility, and comparison of consumer and research-grade activity trackers in insomnia disorder I: in-lab validation against polysomnography. J. Sleep Res. 29, e12931 https://doi.org/10.1111/jsr.12931.
Kessler, R.C., Andrews, G., Mroczek, D., Ustun, B., Wittchen, H.U., 1998. The World Health Organization composite international diagnostic interview short-form (CIDI-SF). Int. J. Methods Psychiatr. Res. 7 (4), 171–185. https://doi.org/10.1002/mpr.47.
Koerber, D., Khan, S., Shamsheri, T., Kirubarajan, A., Mehta, S., 2022. Accuracy of heart rate measurement with wrist-worn wearable devices in various skin tones: a systematic review. J. Racial Ethn. Health Disparities 1-9. https://doi.org/10.1007/s40615-022-01446-9.
Lee, X.K., Chee, N.I.Y.N., Ong, J.L., Teo, T.B., Van, R.E., Lo, J.C., Chee, M.W.L., 2019. Validation of a consumer sleep wearable device with actigraphy and polysomnography in adolescents across sleep opportunity manipulations. J. Clin. Sleep Med. 15, 1337–1346. https://doi.org/10.5664/jcsm.7932.
Leightley, D., Lavelle, G., White, K.M., Sun, S., Matcham, F., Ivan, A., et al., 2021. Investigating the impact of COVID-19 lockdown on adults with a recent history of recurrent major depressive disorder: a multi-Centre study using remote measurement technology. BMC Psychiatr. 21, 435. https://doi.org/10.1186/s12888-021-03434-5.
Lemoine, N.P., 2019. Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. Oikos 128 (7), 912–928. https://doi.org/10.1111/oik.05985.
Mansournia, M.A., Altman, D.G., 2018. Population attributable fraction. BMJ 360, k757. https://doi.org/10.1136/bmj.k757.
Matcham, F., di San, Barattieri, Pietro, C., Bulgari, V., de Girolamo, G., Dobson, R., Eriksson, H., et al., 2019. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. BMC Psychiatr. 19 (1), 72. https://doi.org/10.1186/s12888-019-2049-z.
Matcham, F., Leightley, D., Siddi, S., Lamers, F., White, K.M., Annas, P., et al., 2022. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): recruitment, retention, and data availability in a longitudinal remote measurement study. BMC Psychiatr. 22 (1), 136. https://doi.org/10.1186/s12888-022-03753-1.
Meyer, N., Lok, R., Schmidt, C., Kyle, S.D., McClung, C.A., Cajochen, C., Scheer, F.A.J.L., Jones, M.W., Chellappa, S.L., 2024. The sleep–circadian interface: a window into mental disorders. Proc. Natl. Acad. Sci. 121, e2214756121 https://doi.org/10.1073/pnas.2214756121.
Nelson, B.W., Low, C.A., Jacobson, N., Arean, P., Torous, J., Allen, N.B., 2020. Guidelines for wrist-worn consumer wearable assessment of heart rate in biobehavioral research. NPJ Dig. Med. 3 (90) https://doi.org/10.1038/s41746-020-0297-4.
Novick, J.S., Stewart, J.W., Wisniewski, S.R., Cook, I.A., Manev, R., Nierenberg, A.A., et al., 2005. Clinical and demographic features of atypical depression in outpatients with major depressive disorder: preliminary findings from STAR*D. J. Clin. Psychiatry 66 (8), 1002–1011. https://doi.org/10.4088/jcp.v66n0807.
Oetzmann, C., White, K.M., Ivan, A., Julie, J., Leightley, D., Lavelle, G., Lamers, F., Siddi, S., Annas, P., Garcia, S.A., Haro, J.M., Mohr, D.C., Penninx, B.W.J.H., Simblett, S.K., Wykes, T., Narayan, V.A., Hotopf, M., Matcham, F., 2022. Lessons learned from recruiting into a longitudinal remote measurement study in major depressive disorder. npj Digit. Med. 5, 1–8. https://doi.org/10.1038/s41746-022-00680-z.
Palmier-Claus, J.E., Myin-Germeys, I., Barkus, E., Bentley, L., Udachina, A., Delespaul PAEG, Dunn, G., 2011. Experience sampling research in individuals with mental illness: reflections and guidance. Acta Psychiatr. Scand. 123 (1), 12–20. https://doi.org/10.1111/j.1600-0447.2010.01596.x.
Pandi-Perumal, S.R., Monti, J.M., Burman, D., Karthikeyen, R., BaHammam, A.S., Warren Spence, D., et al., 2020. Clarifying the role of sleep in depression: a narrative review. Psychiatry Res. 291, 113239 https://doi.org/10.1016/j.psychres.2020.113239.
Polhemus, A.M., Novák, J., Ferrao, J., Simblett, S., Radaelli, M., Locatelli, P., et al., 2020. Human-centered design strategies for device selection in mHealth programs: development of a novel framework and case study. JMIR Mhealth Uhealth 8 (5), e16043. https://doi.org/10.2196/16043.
Posternak, M.A., 2003. Biological markers of atypical depression. Harv. Rev. Psychiatry 11 (1), 1–7. https://doi.org/10.1080/10673220303941.
Scott, J., Kallestad, H., Vedaa, O., Sivertsen, B., Etain, B., 2021. Sleep disturbances and first onset of major mental disorders in adolescence and early adulthood: a systematic review and meta-analysis. Sleep Med. Rev. 57, 101429 https://doi.org/10.1016/j.smrv.2021.101429.
Simblett, S., Matcham, F., Siddi, S., Bulgari, V., di San, Barattieri, Pietro, C., Hortas Lopez, J., et al., 2019. Barriers to and facilitators of engagement with mHealth technology for remote measurement and management of depression: qualitative analysis. JMIR Mhealth Uhealth 7 (1), e11325. https://doi.org/10.2196/11325.
Sivula, T., Magnusson, M., Matamoros, A.A., Vehtari, A., 2022. Uncertainty in Bayesian leave-one-out cross-validation based model comparison. http://arxiv.org/abs/2008.10296.
Sjölander, A., Vansteelandt, S., 2019. Frequentist versus Bayesian approaches to multiple testing. Eur. J. Epidemiol. 34 (9), 809–821. https://doi.org/10.1007/s10654-019-00517-2.
Sun, Y., Shi, L., Sun, Y., Shi, J., Lu, L., 2018. The bidirectional relationship between sleep duration and depression in community-dwelling middle-aged and elderly individuals: evidence from a longitudinal study. Sleep Med. 52, 221–229. https://doi.org/10.1016/j.sleep.2018.03.011.
Sun, S., Folarin, A.A., Zhang, Y., Cummins, N., Garcia Dias, R., Stewart, C., et al., 2023. Challenges in the analysis of mobile health data from smartphones and wearable devices to predict depression symptom severity. JMIR 25, e45233. https://doi.org/10.2196/45233.
Tedesco, S., Sica, M., Ancillao, A., Timmons, S., Barton, J., O'Flynn, B., 2019. Validity of the Fitbit Charge2 and the Garmin vivosmart HR+ in free-living environments in an older adult cohort. JMIR Mhealth Uhealth 7 (6), e13084. https://doi.org/10.2196/13084.
Trivedi, M.H., Rush, A.J., Ibrahim, H.M., Biggs, M.M., Suppes, T., Crismon, M.L., et al., 2004. The inventory of depressive symptomatology, clinician rating (IDS-C) and self-report (IDS-SR), and the quick inventory of depressive symptomatology, clinician rating (QIDS-C) and self-report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol. Med. 34 (1), 73–82. https://doi.org/10.1017/s0033291703001107.
Trivedi, M.H., Rush, J.A., Wisniewski, S.R., Nierenberg, A.A., Warden, D., Ritz, L., Fava, M., 2006. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am. J. Psychiatry 163 (1), 28–40. https://doi.org/10.1176/appi.ajp.163.1.28.
Vehtari, A., Gelman, A., Gabry, J., 2016. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27, 1413–1432. https://doi.org/10.1007/s11222-016-9696-4.
Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Burkner, P., Paananen, T., et al.. "loo: efficient leave-one-out cross-validation and WAIC for Bayesian models". R Package Version 2.6.0. https://mc-stan.org/loo/.
Verduijn, J., Verhoeven, J.E., Milaneschi, Y., Schoevers, R.A., Van Hemert, A.M., Beekman, A.T.F., et al., 2017. Reconsidering the prognosis of major depressive disorder across diagnostic boundaries: full recovery is the exception rather than the rule. BMC Med. 15, 215. https://doi.org/10.1186/s12916-017-0972-8.
Vittengl, J.R., Jarrett, R.B., Weitz, E., Hollon, S.D., Twisk, J., Cristea, I., et al., 2016. Divergent outcomes in cognitive-behavioral therapy and pharmacotherapy for adult depression. Am. J. Psychiatry 173 (5), 481–490 (doi:appi.ajp.2015.15040492).
Von Korff, M., Scott, K., Gureje, O., 2009. Global Perspectives on Mental-Physical Comorbidity in the WHO World Mental Health Surveys. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511770531.
Young, M.A., Watel, L.G., Lahmeter, H.W., Eastman, C.I., 1991. The temporal onset of individual symptoms in winter depression: differentiating underlying mechanisms. J. Affect. Disord. 22 (4), 191–197. https://doi.org/10.1016/0165-0327(91)90065-z.
Zhang, Y., Folarin, A.A., Sun, S., Cummins, S., Bendayan, R., Ranjan, Y., et al., 2021. Relationship between major depression symptom severity and sleep collected using a wristband wearable device: multicenter longitudinal observational study. JMIR Mhealth Uhealth 9 (4), e24604. https://doi.org/10.2196/24604.