Unobtrusive inference of circadian rhythms from smartphone data

Loran Knol[1,2,3,4], Mindy K. Ross[3,4], Anisha Nagpal[3], Andrew P. Burns[4], Zachery D. Morrissey[3,5,6,7], Faraz Hussain[3], Tory A. Eisenlohr-Moul[3], Christian F. Beckmann[1,2,8], Alex Leow[3,4,9], Andre F. Marquand[1,2]

[1]Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands

[2]Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen, the Netherlands

[3]Department of Psychiatry, University of Illinois at Chicago, Chicago, IL, USA

[4]Department of Biomedical Engineering, University of Illinois at Chicago, Chicago, IL, USA

[5]Department of Anatomy and Cell Biology, University of Illinois at Chicago, Chicago, IL, USA

[6]Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL, USA

[7]Department of Neurobiology, University of Chicago, Chicago, IL, USA

[8]Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, Oxford, UK

[9]Department of Computer Science, University of Illinois at Chicago, Chicago, IL, USA

Abstract

Circadian rhythms are an integral feature of human cognition and psychopathology but are difficult to measure outside laboratory conditions and over long timescales. Smartphones and other wearables are ubiquitous in daily life and therefore uniquely positioned to measure these rhythms non-invasively and continuously via digital phenotyping. Here, we propose a digital phenotyping framework to quantify circadian rhythms in passive data collected unobtrusively from such devices. First, we use our approach to map the sleep-wake cycle in a clinical, female outpatient sample by predicting sleep duration measured by a dedicated wearable device, the Oura Ring, from smartphone typing dynamics collected through the BiAffect platform. Next, we show that our method captures the desynchronisation between the endogenous circadian rhythm and external clock during international travel by applying it to a year's worth of typing data. Our framework provides a starting point for the analysis of circadian variations in behavioural digital phenotyping data, complementing physiological measurements from wearables.

Introduction

Circadian rhythms are a key feature of human biology and behaviour, and are disrupted in many forms of psychopathology[1], including major depressive disorder[2], schizophrenia[3], and bipolar disorder[4]. For example, a person's preferred waking and sleeping time (chronotype), is associated with depressive and manic symptoms such that an evening chronotype (preferring to go to bed late) confers added risk[5,6]. Such rhythms are also apparent in cognition in that different facets of cognition fluctuate uniquely over time, for example during the day, in a highly individualised manner. They depend on an individual's chronotype and interact with factors such as stress, motivation, and the quality of the previous night's sleep[7–9]. Moreover, long-term disruption of circadian rhythm, for instance due to shift work or frequent exposure to jet lag, impacts both physical health and cognitive performance[10,11].

However, studying these rhythms outside the lab and over long timescales remains challenging with traditional methods. For sleep, polysomnography is considered the gold standard but also the most burdensome[12]. Self-report is less burdensome but subjective and prone to recall bias and attrition[13–15]. Similarly, objectively measuring fluctuations in cognition requires repeated administration of tests, which is logistically challenging and burdensome. Recent advancements in smartphone and wearable technology, however, have allowed researchers to passively measure or approximate different aspects of human behaviour and physiology via digital phenotyping[16]. Wearables such as the Apple Watch, Fitbit, and Oura ring have been used to predict biological sleep features using, e.g., actigraphy[17–25], heart rate variability[24–27], and skin temperature[17,22,26]. Smartphones, however, offer a unique advantage for indexing behaviour, as they are ubiquitous in modern life and most people interact with their phone throughout the day.

Over the last eight years, we have developed and validated a leading smartphone-based platform, BiAffect, for unobtrusive monitoring of behavioural and mental health through keyboard typing dynamics and continuous sensor data.[28–39] BiAffect features an iOS app that replaces the default iPhone keyboard and collects typing speed, autocorrect rate, backspace rate, and phone movement and orientation while typing. These typing dynamics are associated with executive processing speed and show a diurnal pattern in optimal typing performance[32,33]. We have also shown that this approach is sensitive to behavioural fluctuations linked to clinically relevant mental states, predicting symptoms such as depressed mood, mania, and anhedonia[28–30,32,38,39]. Meanwhile, research on other smartphone platforms has shown that measurements such as GPS, screen state, accelerometery, and even typing also predict sleep characteristics[40–44]. This warrants the development of a framework that integrates these measures, offering a comprehensive perspective on behavioural and cognitive fluctuations that complements the physiological insights provided by digital wearables.

Here, we provide a framework to detect and quantify circadian rhythms unobtrusively based on smartphone behaviour – leveraging the smartphone's capability to passively index cognition over extended timescales – and other data. This framework is designed to handle a large number of diverse data types from various devices and manage missing data, which is a key challenge in digital phenotyping[45,46]. Our method provides a fully automated approach to detect circadian rhythms by creating a low-rank approximation of a collection of digital phenotyping modalities using graph-regularised singular value decomposition[47], preserving the data's inherent temporal structure. Our approach is

predicated on the assumption that most people have relatively stable 24-hour rhythms, allowing periods of missing data to be imputed by referencing adjacent points in time and corresponding times on other days. This method results in a quantitative, time-resolved measure of device activity that can be used to infer circadian rhythms. An advantage of this approach is that no assumptions are required about when people are most active during the day, allowing for unrestricted chronotypes and non-contiguous sleep episodes.

As a first step, we use this approach to map the rest-activity cycle non-invasively and unobtrusively over extended periods of time, predicting sleep duration from smartphone typing behaviour in a clinical sample. Next, we demonstrate the utility of our method for understanding the desynchronisation between endogenous circadian processes and the external clock by analysing the circadian phase of a year's worth of BiAffect data of author AL, during which she frequently travelled internationally. Our method provides a starting point for the analysis of circadian variations in behavioural patterns collected from dense digital phenotyping data.

Results

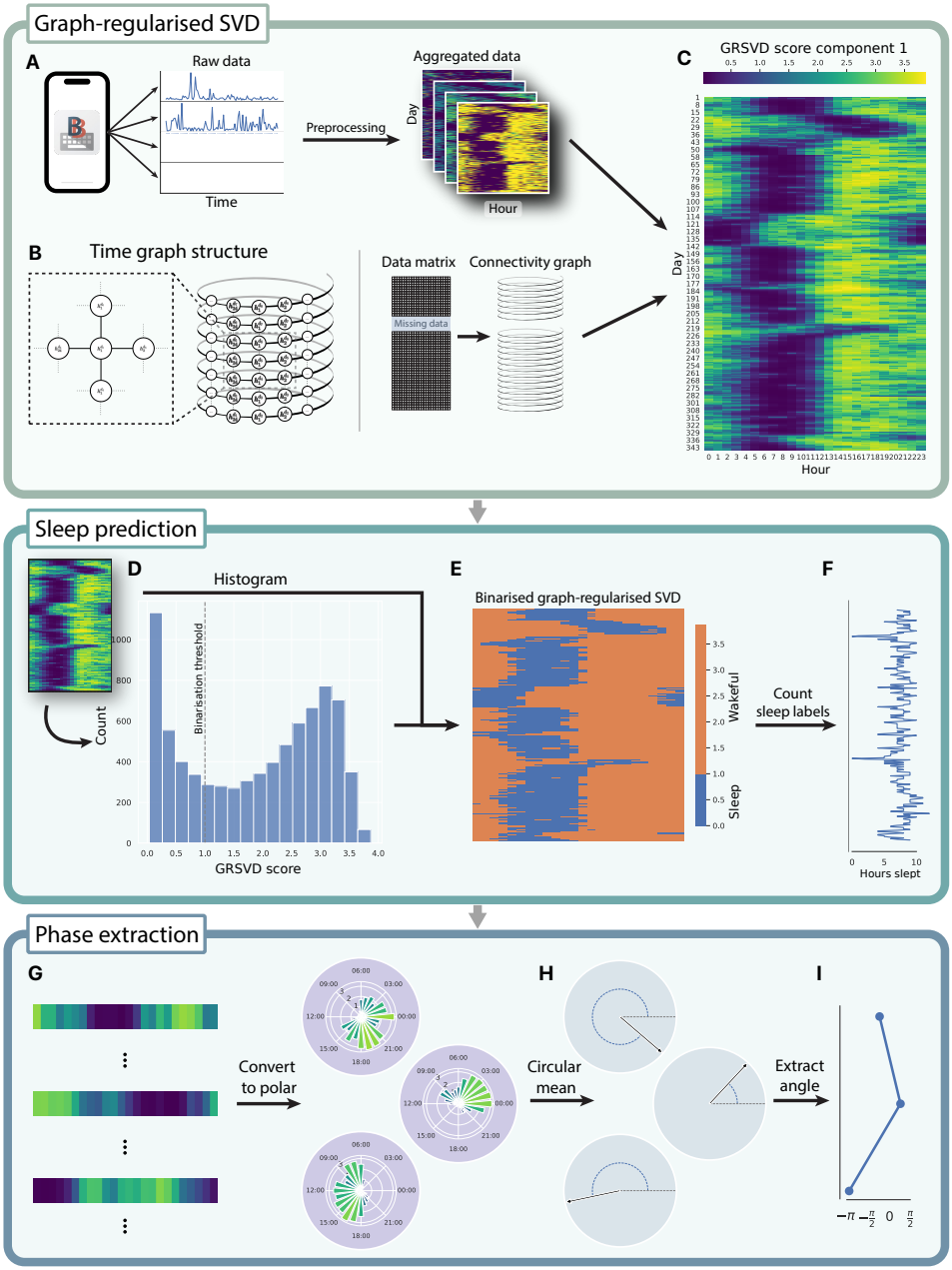

Figure 1: Overview of the graph-regularised SVD. A *Keyboard typing dynamics are collected by the BiAffect app, preprocessed, and aggregated into hourly bins. For every typing modality, the aggregates are arranged into day-by-hour matrices.* ***B*** *Graph of the time structure underlying the hourly bins. Every hour* $\boldsymbol{h}_i^{d_j}$ *is connected to the previous hour* $\boldsymbol{h}_{i-1}^{d_j}$ *and subsequent hour* $\boldsymbol{h}_{i+1}^{d_j}$*. This also means that the final hour of a given day is connected to the first hour of the next day. Similarly, every hour is also connected to the same hour on the previous day* $\boldsymbol{h}_i^{d_{j-1}}$ *and the one on the subsequent day* $\boldsymbol{h}_i^{d_{j+1}}$*. Conceptually, this leads to a graph that resembles a spiral with crosslinks. Missing days of data result in a disjoint graph, preventing information flow to non-adjacent nodes.* ***C*** *The data matrices and time graph are fed into the graph-regularised SVD. This figure shows the first component of the decomposition, where the matrix values can be interpreted as a general measure of typing activity.* ***D*** *Histogram of the matrix values given in C. The dashed line gives the binarisation threshold.* ***E*** *The binarised typing activity matrix. Cells with a value below the binarisation threshold are classified as sleep, and everything else is considered wakeful.* ***F*** *Sleep labels are counted for every day, resulting in a daily sleep duration prediction.* ***G*** *Individual rows of the GRSVD matrix in C can be plotted as a polar bar plot, where the angle is given by the hour in UTC and the height of the bar corresponds to the GRSVD score.* ***H*** *The polar GRSVD scores are summarised by calculating a weighted version of the circular mean.* ***I*** *The angles of the circular mean vectors are extracted, giving the daily phase of the typing activity pattern. GRSVD = Graph-regularised singular value decomposition.*

We sought to develop a framework for the analysis of circadian patterns in digital phenotyping data that would 1) be able to handle the high-dimensional, multimodal nature of such data, 2) respect their underlying temporal structure, and 3) have a certain degree of robustness to missingness. Our solution is based on graph-regularised singular value decomposition (GRSVD) as shown in Figure 1 (see Methods for details). Here, we focus on smartphone typing dynamics as recorded through BiAffect because of their capability to capture diurnal variation as well as BiAffect's minimal impact on smartphone battery life (as opposed to continuously measuring accelerometery, for instance)[32]. It should be noted that this framework is applicable to digital phenotyping data more broadly, however. Once the data have been collected, they need to be preprocessed, after which each modality is aggregated into regular bins – in our case, hours – and arranged into day-by-bin matrices (Figure 1A).

To encode the temporal structure of the data, we construct a graph that specifies which cells in a data matrix are neighbours due to the underlying temporal structure (Figure 1B). More specifically, every hour is connected to the preceding and subsequent hour and to its corresponding hours on the preceding and subsequent day. These connections encode our assumption that a person has a sleep-wake rhythm that slowly fluctuates during the day and is approximately constant across days, enabling the GRSVD to selectively impute missing data by borrowing information from preceding and subsequent hours as well as preceding and subsequent days.

The GRSVD takes both the data matrices and the time graph to produce a dimensionality reduction in which the resulting components have effectively been smoothed based on the graph connectivity (Figure 1C). The amount of regularisation (smoothing) depends on a hyperparameter, which has been optimised by hand (Supplementary Figure 6). The first component reflects the dominant mode of variation across all input modalities and reveals distinct sections of activity and inactivity that capture the user's circadian activity patterns. Other components are shown in Supplementary Figure 11.

While this decomposition provides a promising view on the smartphone activity cycle, it requires further processing to extract relevant features. The sections of inactivity displayed by the SVD matrix prompted us to design an algorithm that could predict sleep duration from smartphone data[48]. First, we determined an appropriate threshold to categorise every hour as either 'sleeping' or 'wakeful' based on the bimodal histogram of the GRSVD matrix (Figure 1D), ensuring the threshold was placed somewhere in the trough between the two peaks. Then, we binarised the matrix based on this threshold (Figure 1E) and, finally, made a daily count of the number of hours with a 'sleep' label to arrive at a prediction for sleep duration (Figure 1F).

Besides the duration of inactivity, the SVD matrix also indicates its on- and offset. This timing is important for the calculation of the phase of the circadian rhythm, which can be influenced by international travel and shift work and is associated with mood disorders[4]. This further allows us to avoid making prior assumptions about someone's chronotype, which provides a major advantage over some of the prior research on smartphones and sleep[40,44]. We quantified the rhythm phase by converting each day (i.e., row) of the GRSVD matrix to polar coordinates, using the GRSVD score as the radius and the hour of day in a specified common reference as the angle (here Coordinated Universal Time (UTC); Figure 1G). From the resulting vectors, we calculated their circular mean, weighing each vector by its magnitude (Figure 1H; Equation 10). These can be interpreted as an indication of

when peak activity occurred during a day. The magnitude gives the strength of the rhythm in that a greater contrast between night- and daytime activity will lead to larger magnitudes, and the angle gives the phase of the daily activity pattern (Figure 1I).

Next, we validated our techniques using 1) a clinical sample with objective sleep duration metrics and 2) a longitudinal sample of international travel.

Smartphone keyboard data predicts sleep duration in a clinical sample

To validate our sleep classification method, we applied it to data from the CLEAR-3 trial, a randomised controlled crossover trial investigating a hormonal intervention to prevent perimenstrual exacerbation of suicidal ideation in people assigned female sex at birth (NCT03498313). For validation, we only used data from the trial's baseline period, when there was no clinical intervention. The participants were encouraged to use BiAffect throughout the study and self-reported on their sleep duration using an hourly Likert scale, with a maximal response of 10+ hours of sleep. A subset also wore Oura Rings (Oura Health Ltd, Oulu, Finland), a digital wearable (smart ring) that estimates sleep duration using accelerometery, photoplethysmography (measuring blood volume changes in the skin using light), and body temperature. These are regarded as good proxies of sleep quality and have been shown to yield 89% to 96% accuracy in detecting wakefulness and sleep as validated against polysomnography[17,27,49]. The version of the Oura Ring used in this study (Generation 2) only recorded bedtime sleep episodes, excluding daytime naps. We generated typing-based estimates for 40 participants with these smart rings (numbers of participants before filtering and merging of data streams are given in Supplementary Figure 1). Demographics are shown in Supplementary Table 1.

To generate features, we extracted hourly aggregates of four BiAffect modalities: 1) Median alphanumeric inter-key delay, which is an inverse measure of typing speed, 2) the total number of key presses, 3) upright typing rate, which measures the proportion of typing sessions in which the phone was oriented upwards, and 4) movement rate, which measures the proportion of typing sessions in which the phone was non-stationary. After filtering to exclude days with too many missing data (see Methods), we calculated the GRSVD of these modalities for every participant (Supplementary Figure 2). We used a threshold of 1.0 to binarise all GRSVD matrices based on visual inspection of their histograms (Figure 1D).

This yielded sleep duration measurements from BiAffect (using GRSVD), the smart ring, and self-report. The BiAffect distribution showed a larger tail than the ring and the (truncated) self-report distributions (Figure 2A). Otherwise, the distributions matched in shape, showing their mode around seven hours of sleep. Linear mixed-effects models showed that our BiAffect-derived estimate of sleep duration significantly predicted ring-derived sleep duration ($n = 953$, $\beta = 0.69$, $p < 0.0001$; Figure 2B)[50], although it showed a tendency for overestimation. Results for different regularisation levels and predictions using self-report data are given in Supplementary Figures 7-10 and Supplementary Table 4. Overall, these findings show that our framework mapped the rest-activity cycle by significantly predicting sleep duration from typing data.

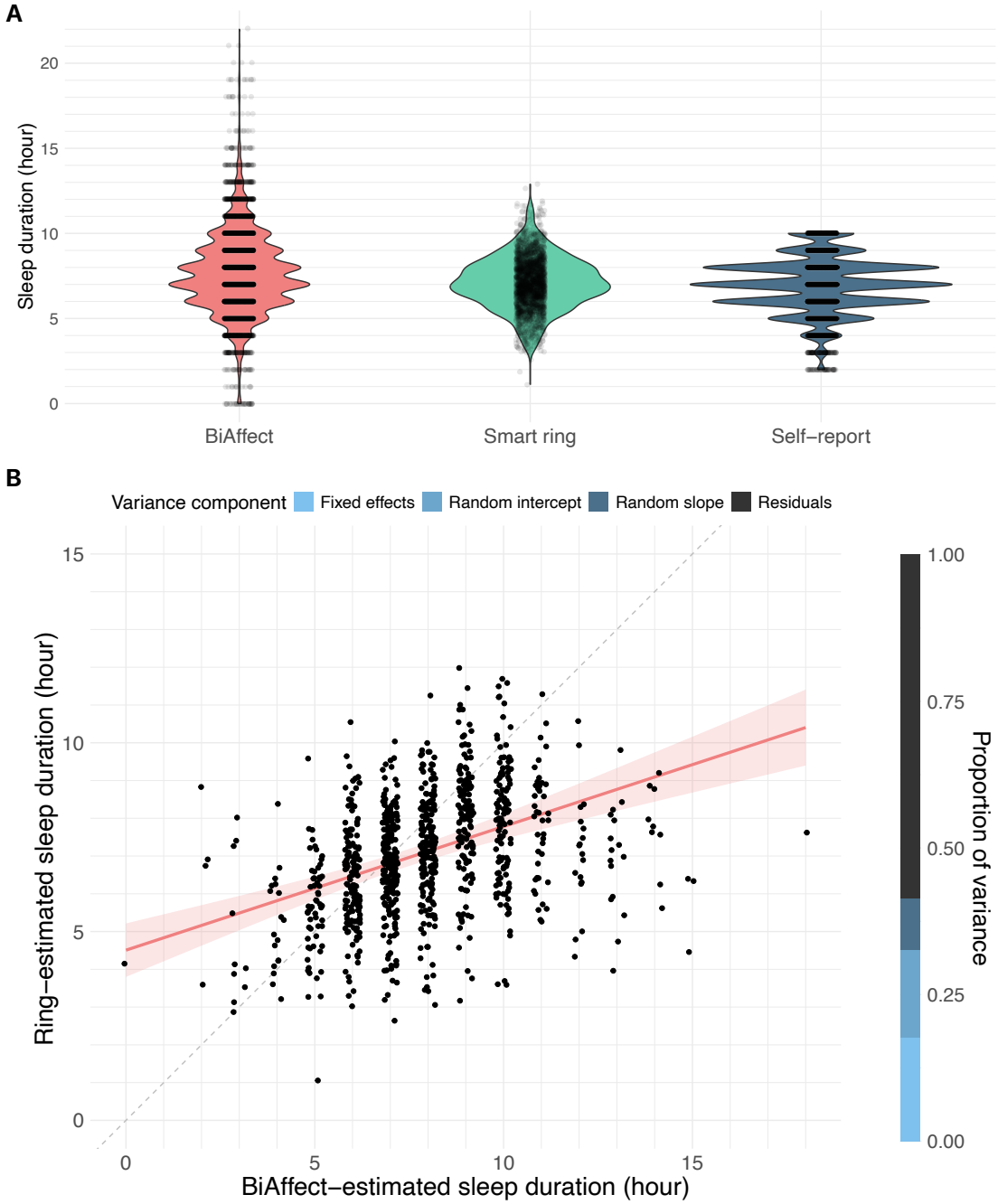

Figure 2: Validation of the graph-regularised SVD method on the CLEAR-3 dataset. A *Violin plots of estimated sleep duration for BiAffect (total n = 6909), the smart ring (n = 2283), and self-report (n=7970). Data points are overlaid to show the data type (continuous or discrete). The typing-derived sleep estimates showed a larger right-skew than the other distributions. These extreme data points corresponded to days for which small amounts of data were collected, possibly due to underutilisation of the phone. In addition, the ring data only concern sleep duration during bedtime, while BiAffect also measures daytime naps if they are present. BiAffect and the ring show a significant Spearman's rank correlation* $r_s = 0.40$ *(n = 953,* $p < 0.0001$*).* ***B*** *Mixed model predictions of ring-derived sleep duration. Data are plotted as jittered points on top of the regression line. The dashed line is the y=x diagonal. The shaded band represents the 95% CI. The horizontal bar graph indicates which proportion of the total variance is explained by which part of the model. The fixed effects explained about 17.8% of the total variance. The random intercepts and slopes explained roughly 14.9% and 8.8% of the variation, respectively. 58.6% of the variance is unexplained.*

Smartphone keyboard data tracks circadian phase

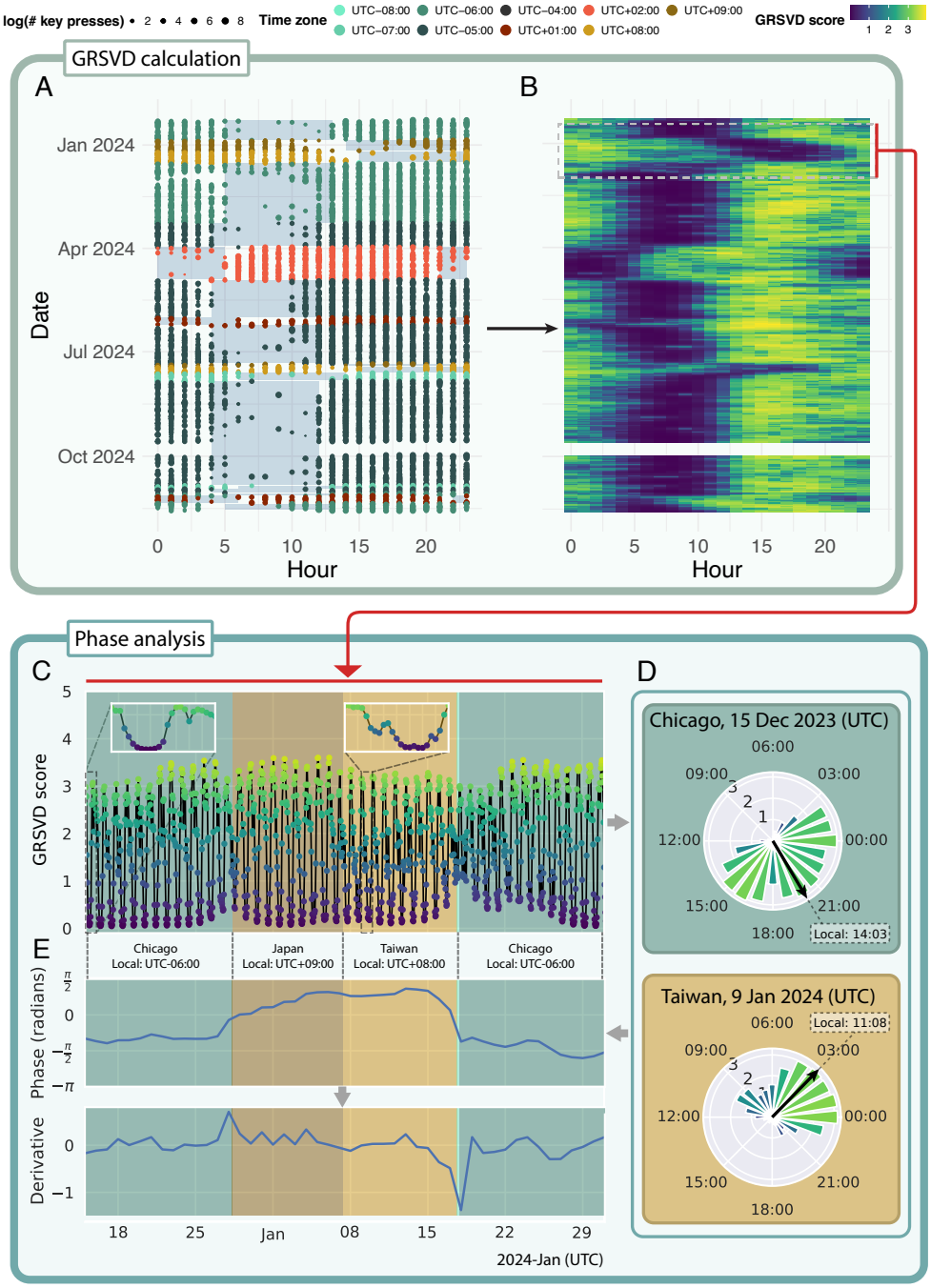

Figure 3: A year's worth of author AL's typing data. A *Overview of the data. Collection started on December 10, 2023, resulting in a nearly continuous data stream from that moment until we pulled the data on November 18, 2024, except for a brief interruption at the end of September 2024 due to technical issues.* ***AL*** *made seventeen transitions across nine time zones. Typing sessions are aggregated into hours and displayed as a matrix of points such that every row corresponds to a day in UTC. Point size and colour correspond to the log key press count and time zone of the phone, respectively. The shaded rectangles represent AL's typical bedtime, 11pm until 7am, in the local time zone.* ***B*** *GRSVD matrix of AL's typing data.* ***C*** *GRSVD scores plotted over time (in UTC) rather than in matrix form. Dates range from December 15, 2023 to January 31, 2024, corresponding to AL's trip from Chicago to Japan and Taiwan. Shaded regions indicate the time zone. Each inset shows a time course of an individual day.* ***D*** *Polar bar plots of the inset time courses shown in C. Bar height represents GRSVD score. Time is shown in UTC. Arrows show angle (not magnitude) of the weighted circular mean of the polar bars and represent the phase delay of the daily rhythm.* ***E*** *Line plots of the daily phase delay and its derivative over time. UTC = Coordinated Universal Time; GRSVD = Graph-regularised singular value decomposition.*

Next, we demonstrate that our approach can automatically detect phase shifts in circadian behaviour. To achieve this, we applied our phase extraction method to a longitudinal sample consisting of a year's worth of BiAffect data from author AL, which contained periods of global travel (Figure 3A). Modality extraction and aggregation proceeded in the same way as for CLEAR-3. Application of the GRSVD exposes the shifts in AL's circadian rhythm as she crosses time zones (Figure 3B). When she crossed ten time zones (travelling from the US to Asia) around January 1st, for instance, her sleep schedule slowly delayed to match her typical sleeping times in the local time zone (shown by the dark blot in Figure 3B shifting to the right). To investigate how the regularity of her circadian rhythm evolved as she travelled, we plotted the GRSVD scores over the time course of this trip (Figure 3C). Every day in this view corresponds to a single oscillation of GRSVD score. The oscillations are initially regular but become clearly disturbed around large time zone transitions with a visible refractory period.

The insets in Figure 3C, which respectively show GRSVD scores of a single day in UTC-06:00 and UTC+08:00, indicate how the circadian rhythm changes with settling into the new time zone. Converting the data of these days to polar form and calculating their circular means gives the phase of the activity patterns (Figure 3D). Figure 3E shows this phase over the entire trip along with its derivative, which quantifies the speed of adaptation to the new time zone (see Supplementary Figure 12 for these measures and the circular mean magnitude over a longer period). We point out that the east- and westwards adaptations did not proceed in the same fashion. Adaptation to UTC+09:00 and UTC+08:00 was gradual and seemed to be complete halfway through the trip. In contrast, when returning to UTC-06:00, the phase snapped back quickly but remained unstable, resulting in overshoot around January 29 (which is also evident from Figure 3B).

Discussion

We proposed a novel framework for the analysis of circadian fluctuations in unobtrusively collected digital phenotyping data. Our method transforms a collection of data modalities to a representation that exposes their principal components through graph-regularised SVD, respecting inherent time dynamics of the data, thus allowing selective imputation[48]. We reported two downstream analyses that quantified different properties of the user's sleep-wake cycle. First, we proposed a method for predicting sleep duration, which we validated with smart ring data in a clinical, psychiatric female outpatient sample, which used the BiAffect keyboard throughout the study. Our BiAffect-derived sleep duration estimates significantly predicted total sleep time as measured by the smart ring. Second, we introduced a method that extracts the phase of the daily device activity pattern and allows comparisons over extended periods of time. We applied this method to a year's worth of author AL's typing data, showing how it can reveal circadian patterns which are not immediately evident from the raw smartphone data.

Despite significantly predicting wearable-derived sleep duration, our typing-based estimates tended to overestimate the actual number of hours slept. This is not surprising, as the most likely cause of this overestimation is that many people do not continuously use their phone while awake and are certain not to use their phone while sleeping. Indeed, if a person mostly uses their phone for tapping and scrolling while only typing occasionally, other smartphone metrics might lead to more accurate sleep predictions, although it should be noted that this tendency for overestimation is displayed by wearables in general[51]. We further emphasise that our framework provides an alternative and complementary perspective on circadian rhythms by considering behaviour and cognition, and should by no means be regarded as a replacement for wearables like the current smart ring, which primarily focus on physiology. Future research on mental health and digital phenotyping will most likely benefit from taking both perspectives into account.

We chose to demonstrate our framework with typing dynamics because of the previously demonstrated relationship between time-of-day and typing speed and the minimal impact that the BiAffect app has on battery life[32]. Other modalities would also have been suitable, however, as shown by several studies using ambient sound and light, GPS, screen state, or accelerometery to predict sleep[40–43]. Most of them used small sample sizes and self-report for validation, but recent studies feature larger sample sizes[52,53]. A wide variety of machine learning models has been used in these studies, including Bayesian networks[43], decision trees[42,43], linear mixed-effects regression[44,53], Bayesian models[52], and rule-based classifiers[40]. Only a handful of these studies focus on the circadian rhythm more generally by, for instance, identifying participant chronotypes[40], and none consider the effects of international travel.

Our computational framework allows the analysis of the desynchronisation of endogenous circadian rhythms from the external clock, which we demonstrated with a year of iPhone data from author AL, leveraging her frequent time zone transitions during that period. More specifically, we showed that time zone transitions resulted in more irregular rhythms and phase shifts of the sleep-wake cycle, which we were able to quantify.

Regularity in our circadian rhythms is important for our health. Shift workers whose schedules misalign with their chronotype (e.g., working an early morning shift as an evening person) show impaired cognition[54], attenuated melatonin fluctuations, more sleep disturbance, and larger social jet lag[55,56]. The impact of these changes go beyond poor sleep as melatonin has an anti-metastatic effect, potentially explaining the higher rates of breast cancer among night-shift workers[11]. Similar results have been found for airline cabin crew[57,58]. Although completely eliminating shift work may not be possible in present-day society, maintaining a real-time index of the phase and strength of someone's circadian rhythm could be helpful in monitoring its long-term disruption and facilitate intervention when those disruptions become too severe. We plan to add this feature to the BiAffect dashboard to allow users to monitor their own circadian rhythm.

One potential improvement to our framework would be to include more extensive capabilities to differentiate between different kinds of missing data, such as technical issues versus a lack of phone usage. Another option would be to extend our framework to group-level decompositions, such as in the work of Aledavood et al., who applied non-negative matrix factorisation to smartphone data to identify participant chronotypes[59]. Aside from such optimisations, future research should also investigate how smartphone rhythms in general respond to psychopathological disturbances, such as the onset of mood episodes in bipolar disorder. Once these dynamics have been determined, these fluctuations may serve as digital markers of mental wellbeing and be incorporated into an early signalling system that could warn doctor and patient about deviations from someone's baseline rhythms.

We conclude that our GRSVD was successful in uncovering circadian patterns in digital phenotyping data and has several appealing features. First, our method makes few assumptions about the participant's chronotype or the contiguity of their sleep, which contrasts with some of the existing methods[40,44]. Other advantages include its capability to incorporate the continuous and cyclical nature of time through a graph, allowing the method to selectively impute missing data using surrounding time points, as well as the ability to deal with a large number of modalities. In this study, we only included four modalities originating from one device, but this could easily be expanded to include other data types from a variety of platforms, with the additional possibility to retain more than just one dimension of the GRSVD, which could provide richer characterisation of circadian dynamics.

In sum, this approach provides a promising new method to index circadian rhythms unobtrusively and over long timescales, providing a powerful tool for future clinical studies and a principled method for individuals to monitor their own circadian dynamics.

Methods

Preprocessing

Preprocessing of BiAffect data largely followed the same procedure as in previous works[37,38], briefly summarised here. To limit the size of the typing data matrices, we aggregated into hourly bins using several methods, yielding different features for analysis. First, we counted the total number of key presses for every hour. To constrain the range of key press counts and ensure positivity, we incremented them by one and took the natural log. We then calculated the inter-key delays (IKDs) between all alphanumeric key presses within a single typing session, defined as starting at the first key press and terminating when the participant either closed the keyboard or did not enter any key press for six seconds. Consequently, we took the median of these IKDs for every hour to derive an inverse measure of typing speed[37,32].

Accelerometer data were collected only when a typing session was active. Data were sampled at 10 Hz and filtered with a second-order bidirectional Butterworth filter with a cut-off of 4 Hz to remove high-frequency noise. We then classified typing sessions as moving or 'active' if more than 8% of the accelerometer samples had a magnitude lower than 0.95 or higher than 1.05 (a magnitude of 1 reflects the gravitational pull of the earth). To classify sessions as upright or lying down, we considered the median x- and z-values of the accelerometer samples: We would mark a session as upright if the median x-value was in-between -0.2 and 0.2 (inclusive) and the median z-value was below 0.1[37]. Otherwise, the session was considered as lying down. Finally, these Boolean session classifications were averaged per hour, giving rates of moving and upright sessions.

Since our method is contingent on the presence of enough data to distinguish between waking and sleeping hours, we applied additional processing steps before applying the graph-regularised SVD. First, we right-shifted all timestamps with +12 hours to bring the nights to the centre of our 24-hour periods, which does not change the outcome of the graph-regularised SVD but instead facilitates the calculation of the amount of time slept during those nights (see Figure 4). From this point on, a 'day' is defined as the time period running from 12pm on one day to 11:59am on the next day. After that, we required participants to have at least, on average, 1) 20% of their daily hours filled with some data,

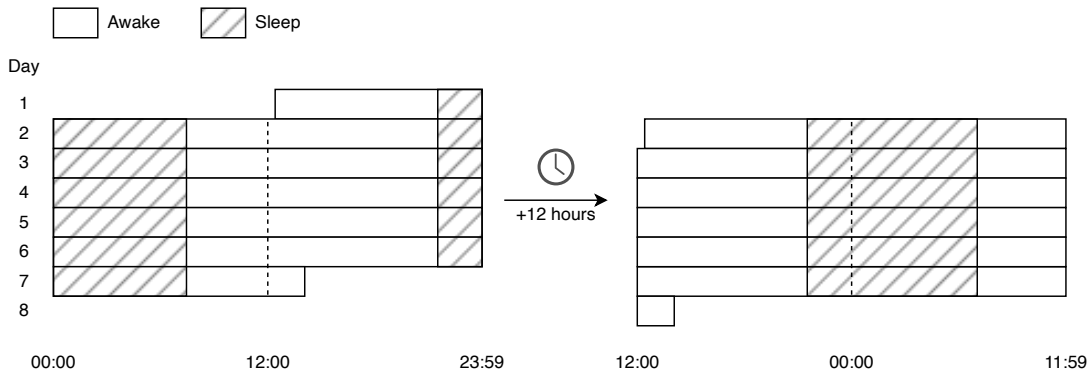

Figure 4: Schematic overview of the time shift applied during preprocessing. *Adding +12 hours to every data timestamp will ensure that every 24-hour period, represented as a row in our data matrix, has a complete night. Otherwise, every night could be dispersed over two rows. While this does not affect the outcome of the graph-regularised SVD, it does make it easier to count the number of hours slept in the time windows that the smart ring uses.*

2) 50 key presses per hour, and 3) nine days of data. They would be excluded from the data set if they did not meet these criteria. For remaining participants, we required at least 33% of every day to have data (results for other thresholds are given in Supplementary Figures 4-5 and Supplementary Table 3).

All typing-derived features were scaled to unit variance (but not mean-centred) within participants, separately for the training and testing sets (described below).

This study used the Oura Ring Generation 2 to measure sleep duration. Oura ring total sleep time was derived according to proprietary algorithms. The Oura ring uses measures such as light exposure, actigraphy, and estimates of someone's circadian rhythm to determine their bed- and waketime and only measures sleep episodes during the night[17]. This means that daytime naps are not reflected in the total sleep time.

Graph-regularised SVD

The goal of the decomposition or factorisation of a matrix is often to use the decomposition to find a lower-dimensional approximation to the full matrix. In the context of this study, we aimed to employ such a factorisation to reduce a set of typing modalities to a single measure of typing activity. A well-known factorisation method is singular value decomposition (SVD):

$$X = U\Sigma V^{\top}, \tag{1}$$

where X is an $m \times n$ matrix, U is an $m \times m$ orthogonal matrix, Σ is an $m \times n$ diagonal matrix containing the singular values of X, and V is an $n \times n$ orthogonal matrix containing the eigenvectors of $X^{\top}X$[60]. By picking only those vectors in U and V corresponding to the largest singular values, this is equivalent to principal component analysis, thereby reducing the dimensionality of the data.

Graph-regularised matrix decomposition attempts to do this decomposition whilst respecting the geometrical structure of the data[61]. The data geometry is represented as a graph that is encoded in a (symmetric) adjacency matrix A. Its entries $a_{ij} \in \{0, 1\}$ give the connection strength between nodes i and j. In this work, we restrict a_{ij} to be binary but they could also be continuous such as in the case of a weighted graph. To incorporate A into the decomposition, its corresponding graph Laplacian is calculated as $L = D - A$, where D is a diagonal matrix containing the sums of the columns (or rows) of A. We then seek a decomposition of X by solving the following problem[61]:

$$\min_{H,W}\left(\|X - HW\|^2 + \lambda Tr(W^{\top}LW)\right), \tag{2}$$

where H is an $m \times k$ matrix, W is a $k \times n$ matrix, and λ is a regularisation parameter controlling the smoothness of the decomposition $HW \approx X$. k determines the rank of the approximation, i.e., it represents the dimensionality that we reduce our data to.

Minimising the objective in Equation (2) starts by computing an SVD[47]:

$$XR^{\top^{-1}} = U\Sigma V^{\top}, \tag{3}$$

where $RR^{\top} = I + \lambda L$ is an LU decomposition, in which I is the identity matrix. We can then approximate the SVD by taking only the k largest singular values and the corresponding vectors in U and V:

$$XR^{\top^{-1}} \approx \widetilde{U}\widetilde{\Sigma}\widetilde{V}^{\top}. \tag{4}$$

Consequently, by setting

$$\widehat{H} \coloneqq \widetilde{U}, \tag{5}$$

$$\widehat{W} \coloneqq \widetilde{\Sigma}\widetilde{V}^{\top}R^{-1}, \tag{6}$$

the objective in Equation (2) will be minimised[62].

In this study, the data matrix X contains $m = 4$ typing features derived from a participant's keyboard usage, aggregated into hourly bins. These bins form the nodes of the time graph and are connected to 1) their preceding and subsequent neighbours and 2) the corresponding hours of the preceding and subsequent day. In the case of missing data, we simply remove nodes corresponding to unrecorded days from the graph. This leads to disjoint graphs but poses no problem for the analysis. Since every hour is a node, we have $n = 24 \cdot d$ nodes, where d is the number of days we have recorded for the participant. Our aim is to reduce the $4 \times 24d$ matrix of typing features X to a $1 \times 24d$ vector which gives us a general measure of typing activity for every hour of every day. This can easily be done by setting $k = 1$ and taking the right-hand factor of the graph-regularised decomposition, $\widehat{W}$, as our activity vector.

The left factor $\widetilde{U}$ forms a basis transformation that can be applied to new data X^* and the associated Laplacian L^*. This allows us to 'train' the decomposition on a subset of a participant's data and then apply or 'test' it on the remainder. First we rotate the test data into the same space as $\widetilde{\Sigma}\widetilde{V}^{\top}$:

$$B = \widetilde{U}^{\top}X^*R^{*^{\top^{-1}}}, \tag{7}$$

where $R^*R^{*^{\top}} = I + \lambda L^*$. We then set, analogously to how we constructed $\widehat{W}$:

$$\widehat{W}^* \coloneqq BR^{*^{-1}}. \tag{8}$$

Since $\widetilde{U}$ only contains four values in our case, it does not require much data to train. We therefore decided to use 10% of a participant's data for training. The remaining 90% was used to predict the number of hours slept, with the exclusion of the first three days after the training period to prevent information from the training set leaking into the test set through temporal autocorrelation. (For the GRSVD of author AL's data, we instead used a 5/95 train-test split due to the large number of available data. In addition, we trained on the tail (the last 5%) rather than the head of that data set.) Participants that did not have enough days of data to make this train-test split (i.e., less than 10 days of data) were excluded from the analysis.

Sleep duration prediction

To predict sleep, the GRSVD scores were binarised into two categories: All values below some threshold were classified as sleep, and the rest was classified as wakeful. This threshold was determined from visual inspection of the SVD histograms and positioned to be in-between the two peaks of the bimodal distribution. From that point, the number of hours classified as sleeping were counted to arrive at a prediction for sleep duration. For validation, we used nightly aggregates of the total sleep time as measured by the

smart ring and predicted them with our own sleep duration estimates using linear mixed-effects models.

The mixed-effects models included a random intercept b_{0i} and a random slope for typing-derived sleep estimates b_{1i} per participant i:

$$y_{ij} = \beta_0 + b_{0i} + \beta_1 x_{ij} + b_{1i} x_{ij} + \epsilon_{ij}, \tag{9}$$

where y_{ij}, x_{ij}, and ϵ_{ij} respectively represent the ring-derived total sleep time, typing-derived sleep estimate, and error for participant i and day j, while β_0 and β_1 give, respectively, the grand mean and fixed effect of our estimate across participants. BiAffect-derived sleep duration was standardised (M = 0, SD = 1) across all observations. Model assumptions were checked using visual inspection of the QQ plots of the residuals and random effects. No gross violations of these assumptions were detected.

Phase extraction

We converted every entry of AL's GRSVD matrix to polar coordinates (r, θ), where the radius r corresponds to the entry's SVD value and the angle $\theta = \frac{h}{24} \cdot 2\pi$ corresponds to the hour of day h rescaled to $[0, 2\pi)$. We then calculated a weighted version of the circular mean for every row (day) of the SVD matrix:

$$\bar{\rho} = \frac{1}{N} \sum_{n=1}^{N} r_n \cdot e^{i\theta_n} = \frac{1}{N} \sum_{n=1}^{N} r_n (cos\, \theta_n + i\, sin\, \theta_n)\,, \tag{10}$$

where $N = 24$. The angle of this circular mean can then be calculated by:

$$\bar{\theta} = atan2\left(\frac{1}{N} \sum_{n=1}^{N} r_n\, sin\, \theta_n\,, \frac{1}{N} \sum_{n=1}^{N} r_n\, cos\, \theta_n \right). \tag{11}$$

Similarly, for its radius:

$$\bar{r} = \sqrt{\left(\frac{1}{N} \sum_{n=1}^{N} r_n\, cos\, \theta_n \right)^2 + \left(\frac{1}{N} \sum_{n=1}^{N} r_n\, sin\, \theta_n \right)^2}\,. \tag{12}$$

Study design

We used data from a randomised crossover trial (NCT04112368) that investigated the administration of estradiol (E2) and progesterone (P4) to prevent perimenstrual exacerbation of suicidal ideation due to the natural withdrawal of E2 and P4 in individuals assigned the female sex at birth (AFAB). The study was approved by the UIC Institutional Review Board.

This study focused only on the baseline data of the trial (i.e., prior to the intervention and placebo), which consisted of at least one menstrual cycle. Participants self-reported on last-night sleep duration, and some of them wore Oura rings which collected total sleep duration through physiological measurements[17]. Finally, participants were encouraged to use the BiAffect keyboard as their primary keyboard for the duration of the study.

Recruitment and exclusion criteria

Participants were recruited via social media. They were 18-45 years of age, had menstrual cycles of 25-35 days, did not take hormonal medication and were in the BMI range 18-34.99. All of them reported past-month suicidal ideation and were in stable outpatient psychiatric treatment. Upon completion of the trial, they received up to US$1250. Participants with any serious long-term nonpsychiatric health condition, a history of hospitalisation for mania or psychosis, or any affective or substance use disorder that would pose a high risk to safe participation in the trial were excluded. All participants provided informed consent.

Data availability

Deidentified participant data will be made available on reasonable request to the principal investigator of the CLEAR-3 trial, T.A. Eisenlohr-Moul (temo@uic.edu).

Code availability

All analysis source code is freely available on GitHub: https://github.com/Valkje/graph-svd.

Acknowledgements

This study was funded by the European Research Council (101001118). TAE-M and AN received funding from the National Institute of Mental Health (RF1MH120843 and F30MH138058, respectively). CFB gratefully acknowledges funding from the Wellcome Trust Collaborative Award in Science 215573/Z/19/Z and the Netherlands Organization for Scientific Research Vici Grant No. 17854 and NWO-CAS Grant No. 012-200-013.

Author contributions

LK: Conceptualisation; data curation; software; formal analysis; methodology; visualisation; writing – original draft. MKR: Conceptualisation; methodology; software. AN: Data curation; writing – review & editing. APB: Conceptualisation; methodology; software; writing – review & editing. ZDM: Conceptualisation; methodology; software; writing – review & editing. FH: Software; writing – review & editing. TAE-M: Data curation; funding acquisition; investigation; writing – review & editing. CFB: Funding acquisition; supervision; writing – review & editing. AL: Conceptualisation; funding acquisition; investigation; methodology; supervision; writing – review & editing. AFM: Conceptualisation; funding acquisition; methodology; supervision; writing – review & editing.

Competing interests

CFB is founding director and shareholder of SBG Neuro Ltd. AL is a cofounder of KeyWise AI, previously served on the Medical Board for Buoy Health, and currently serves as a digital psychiatry advisor for Otsuka, USA.

References

1. Logan, R. W. & McClung, C. A. Rhythms of life: circadian disruption and brain disorders across the lifespan. *Nat. Rev. Neurosci.* **20**, 49–65 (2019).
2. Tsuno, N., Besset, A. & Ritchie, K. Sleep and Depression. *J. Clin. Psychiatry* **66**, 19685 (2005).
3. Chouinard, S., Poulin, J., Stip, E. & Godbout, R. Sleep in Untreated Patients With Schizophrenia: A Meta-Analysis. *Schizophr. Bull.* **30**, 957–967 (2004).
4. Gold, A. K. & Sylvia, L. G. The role of sleep in bipolar disorder. *Nat. Sci. Sleep* **8**, 207–214 (2016).
5. Tonon, A. C. *et al.* Sleep and circadian disruption in bipolar disorders: From psychopathology to digital phenotyping in clinical practice. *Psychiatry Clin. Neurosci.* **78**, 654–666 (2024).
6. Lewy, A. J. Depressive disorders may more commonly be related to circadian phase delays rather than advances: Time will tell. *Sleep Med.* **11**, 117–118 (2010).
7. Schmidt, C., Collette, F., Cajochen, C. & Peigneux, P. A time to think: Circadian rhythms in human cognition. *Cogn. Neuropsychol.* **24**, 755–789 (2007).
8. Könen, T., Dirk, J. & Schmiedek, F. Cognitive benefits of last night's sleep: daily variations in children's sleep behavior are related to working memory fluctuations. *J. Child Psychol. Psychiatry* **56**, 171–182 (2015).
9. Neubauer, A. B., Dirk, J. & Schmiedek, F. Momentary working memory performance is coupled with different dimensions of affect for different children: A mixture model analysis of ambulatory assessment data. *Dev. Psychol.* **55**, 754–766 (2019).
10. Marquié, J.-C., Tucker, P., Folkard, S., Gentil, C. & Ansiau, D. Chronic effects of shift work on cognition: findings from the VISAT longitudinal study. *Occup. Environ. Med.* **72**, 258–264 (2015).
11. Hill, S. M. *et al.* Melatonin: an inhibitor of breast cancer. *Endocr. Relat. Cancer* **22**, R183–R204 (2015).
12. Rundo, J. V. & Downey, R. Chapter 25 - Polysomnography. in *Handbook of Clinical Neurology* (eds. Levin, K. H. & Chauvel, P.) vol. 160 381–392 (Elsevier, 2019).
13. Cudney, L. E., Frey, B. N., McCabe, R. E. & Green, S. M. Investigating the relationship between objective measures of sleep and self-report sleep quality in healthy adults: a review. *J. Clin. Sleep Med.* **18**, 927–936.
14. Lauderdale, D. S., Knutson, K. L., Yan, L. L., Liu, K. & Rathouz, P. J. Self-Reported and Measured Sleep Duration: How Similar Are They? *Epidemiology* **19**, 838 (2008).
15. Wrzus, C. & Neubauer, A. B. Ecological Momentary Assessment: A Meta-Analysis on Designs, Samples, and Compliance Across Research Fields. *Assessment* **30**, 825–846 (2023).
16. Torous, J., Kiang, M. V., Lorme, J. & Onnela, J.-P. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. *JMIR Ment. Health* **3**, e5165 (2016).

17. Altini, M. & Kinnunen, H. The Promise of Sleep: A Multi-Sensor Approach for Accurate Sleep Stage Detection Using the Oura Ring. *Sensors* **21**, 4302 (2021).

18. de Zambotti, M., Goldstone, A., Claudatos, S., Colrain, I. M. & Baker, F. C. A validation study of Fitbit Charge 2™ compared with polysomnography in adults. *Chronobiol. Int.* **35**, 465–476 (2018).

19. van Wouwe, N. C., Valk, P. J. L. & Veenstra, B. J. Sleep Monitoring: A Comparison Between Three Wearable Instruments. *Mil. Med.* **176**, 811–816 (2011).

20. Sathyanarayana, A. *et al.* Sleep Quality Prediction From Wearable Data Using Deep Learning. *JMIR MHealth UHealth* **4**, e6562 (2016).

21. Kazemi, K., Azimi, I., Liljeberg, P. & Rahmani, A. M. Can Sleep Quality Attributes be Predicted from Physical Activity in Everyday Settings? in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* 1–5 (2023). doi:10.1109/EMBC40787.2023.10340421.

22. Logacjov, A., Skarpsno, E. S., Kongsvold, A., Bach, K. & Mork, P. J. A Machine Learning Model for Predicting Sleep and Wakefulness Based on Accelerometry, Skin Temperature and Contextual Information. *Nat. Sci. Sleep* **16**, 699–710 (2024).

23. Trujillo, R., Zhang, E., Templeton, J. M. & Poellabauer, C. Predicting long-term sleep deprivation using wearable sensors and health surveys. *Comput. Biol. Med.* **179**, 108749 (2024).

24. Walch, O., Huang, Y., Forger, D. & Goldstein, C. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. *Sleep* **42**, zsz180 (2019).

25. Park, K. M. *et al.* Predicting sleep based on physical activity, light exposure, and Heart rate variability data using wearable devices. *Ann. Med.* **56**, 2405077 (2024).

26. Sadeghi, R., Banerjee, T., Hughes, J. C. & Lawhorne, L. W. Sleep quality prediction in caregivers using physiological signals. *Comput. Biol. Med.* **110**, 276–288 (2019).

27. Miller, D. J., Sargent, C. & Roach, G. D. A Validation of Six Wearable Devices for Estimating Sleep, Heart Rate and Heart Rate Variability in Healthy Adults. *Sensors* **22**, 6317 (2022).

28. Cao, B. *et al.* DeepMood: Modeling Mobile Phone Typing Dynamics for Mood Detection. in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 747–755 (Association for Computing Machinery, New York, NY, USA, 2017). doi:10.1145/3097983.3098086.

29. Stange, J. P. *et al.* Let your fingers do the talking: Passive typing instability predicts future mood outcomes. *Bipolar Disord.* **20**, 285–288 (2018).

30. Zulueta, J. *et al.* Predicting Mood Disturbance Severity with Mobile Phone Keystroke Metadata: A BiAffect Digital Phenotyping Study. *J. Med. Internet Res.* **20**, e9775 (2018).

31. Hussain, F. *et al.* Passive Sensing of Affective and Cognitive Functioning in Mood Disorders by Analyzing Keystroke Kinematics and Speech Dynamics. in *Digital Phenotyping and Mobile Sensing: New*

Developments in Psychoinformatics (eds. Baumeister, H. & Montag, C.) 161–183 (Springer International Publishing, Cham, 2019). doi:10.1007/978-3-030-31620-4_10.

32. Vesel, C. *et al.* Effects of mood and aging on keystroke dynamics metadata and their diurnal patterns in a large open-science sample: A BiAffect iOS study. *J. Am. Med. Inform. Assoc.* **27**, 1007–1018 (2020).

33. Ross, M. K. *et al.* Naturalistic smartphone keyboard typing reflects processing speed and executive function. *Brain Behav.* **11**, e2363 (2021).

34. Zulueta, J. *et al.* The Effects of Bipolar Disorder Risk on a Mobile Phone Keystroke Dynamics Based Biomarker of Brain Age. *Front. Psychiatry* **12**, (2021).

35. Chen, M. H. *et al.* Associations between smartphone keystroke dynamics and cognition in MS. *Digit. Health* **8**, 20552076221143234 (2022).

36. Rashidisabet, H., Ajilore, O., Leow, A. & Demos, A. P. Revisiting power-law estimation with applications to real-world human typing dynamics. *Phys. Stat. Mech. Its Appl.* **599**, 127384 (2022).

37. Ning, E. *et al.* Smartphone-derived Virtual Keyboard Dynamics Coupled with Accelerometer Data as a Window into Understanding Brain Health: Smartphone Keyboard and Accelerometer as Window into Brain Health. in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* 1–15 (ACM, Hamburg Germany, 2023). doi:10.1145/3544548.3580906.

38. Knol, L. *et al.* Smartphone keyboard dynamics predict affect in suicidal ideation. *Npj Digit. Med.* **7**, 1–10 (2024).

39. Liu, Q. *et al.* Digital Phenotypes of Mobile Keyboard Backspace Rates and Their Associations With Symptoms of Mood Disorder: Algorithm Development and Validation. *J. Med. Internet Res.* **26**, e51269 (2024).

40. Abdullah, S., Matthews, M., Murnane, E. L., Gay, G. & Choudhury, T. Towards circadian computing: 'early to bed and early to rise' makes some of us unhealthy and sleep deprived. in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* 673–684 (Association for Computing Machinery, New York, NY, USA, 2014). doi:10.1145/2632048.2632100.

41. Bai, Y., Xu, B., Ma, Y., Sun, G. & Zhao, Y. Will You Have a Good Sleep Tonight? Sleep Quality Prediction with Mobile Phone. in (2012).

42. Hao, T., Xing, G. & Zhou, G. iSleep: unobtrusive sleep quality monitoring using smartphones. in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* 1–14 (Association for Computing Machinery, New York, NY, USA, 2013). doi:10.1145/2517351.2517359.

43. Min, J.-K. *et al.* Toss 'n' turn: smartphone as sleep and sleep quality detector. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 477–486 (Association for Computing Machinery, New York, NY, USA, 2014). doi:10.1145/2556288.2557220.

44. Druijff-van de Woestijne, G. B. *et al.* Behavioural biometrics: Using smartphone keyboard activity as a proxy for rest–activity patterns. *J. Sleep Res.* **30**, e13285 (2021).

45. Currey, D. & Torous, J. Increasing the value of digital phenotyping through reducing missingness: a retrospective review and analysis of prior studies. *BMJ Ment Health* **26**, (2023).

46. Bloom, P. A. *et al.* Identifying factors impacting missingness within smartphone-based research: Implications for intensive longitudinal studies of adolescent suicidal thoughts and behaviors. *J. Psychopathol. Clin. Sci.* No Pagination Specified-No Pagination Specified (2024) doi:10.1037/abn0000930.

47. Vidar, E. A. & Alvindia, S. K. SVD Based Graph Regularized Matrix Factorization. in *Intelligent Data Engineering and Automated Learning – IDEAL 2013* (eds. Yin, H. et al.) 234–241 (Springer, Berlin, Heidelberg, 2013). doi:10.1007/978-3-642-41278-3_29.

48. Ross, M. K. *et al.* Keying into Cognition: Temporal Smoothing of Smartphone Typing Behaviors for Passive Assessment of Processing Speed and Executive Function in Individuals with Mood Disorders. Preprint at https://doi.org/10.21203/rs.3.rs-4849283/v1 (2024).

49. de Zambotti, M., Rosas, L., Colrain, I. M. & Baker, F. C. The Sleep of the Ring: Comparison of the ŌURA Sleep Tracker Against Polysomnography. *Behav. Sleep. Med.* **17**, 124–136 (2019).

50. Pinheiro, J. C. & Bates, D. M. *Mixed-Effects Models in S and S-PLUS*. (Springer-Verlag, New York, 2000). doi:10.1007/b98882.

51. Kolla, B. P., Mansukhani, S. & Mansukhani, M. P. Consumer sleep tracking devices: a review of mechanisms, validity and utility. *Expert Rev. Med. Devices* **13**, 497–506 (2016).

52. Cuttone, A. *et al.* SensibleSleep: A Bayesian Model for Learning Sleep Patterns from Smartphone Events. *PLOS ONE* **12**, e0169901 (2017).

53. Langholm, C., Byun, A. J. S., Mullington, J. & Torous, J. Monitoring sleep using smartphone data in a population of college students. *Npj Ment. Health Res.* **2**, 1–6 (2023).

54. Marquié, J.-C., Tucker, P., Folkard, S., Gentil, C. & Ansiau, D. Chronic effects of shift work on cognition: findings from the VISAT longitudinal study. *Occup. Environ. Med.* **72**, 258–264 (2015).

55. Juda, M., Vetter, C. & Roenneberg, T. Chronotype Modulates Sleep Duration, Sleep Quality, and Social Jet Lag in Shift-Workers. *J. Biol. Rhythms* **28**, 141–151 (2013).

56. Razavi, P. *et al.* Shift Work, Chronotype, and Melatonin Rhythm in Nurses. *Cancer Epidemiol. Biomarkers Prev.* **28**, 1177–1186 (2019).

57. Cho, K., Ennaceur, A., Cole, J. C. & Suh, C. K. Chronic Jet Lag Produces Cognitive Deficits. *J. Neurosci.* **20**, RC66 (1-5) (2000).

58. Rafnsson, V., Tulinius, H., Jónasson, J. G. & Hrafnkelsson, J. Risk of breast cancer in female flight attendants: a population-based study (Iceland). *Cancer Causes Control* **12**, 95–101 (2001).

59. Aledavood, T., Kivimäki, I., Lehmann, S. & Saramäki, J. Quantifying daily rhythms with non-negative matrix factorization applied to mobile phone data. *Sci. Rep.* **12**, 5544 (2022).

60. Shlens, J. A Tutorial on Principal Component Analysis. Preprint at https://doi.org/10.48550/arXiv.1404.1100 (2014).

61. Cai, D., He, X., Han, J. & Huang, T. S. Graph Regularized Nonnegative Matrix Factorization for Data Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**, 1548–1560 (2011).

62. Donelli, P. New Functional Data Analysis Methods, with Applications to Spatial Transcriptomics and Neuroimaging. (2023).