The two fundamental shapes of sleep heart rate dynamics and their connection to mental health

Mikaela Irene Fudolig,[1,2,3,*] Laura S. P. Bloomfield,[4,1,2,3] Matthew Price,[5,1] Yoshi M. Bird,[1,3] Johanna E. Hidalgo,[5] Julia Kim,[6] Jordan Llorin,[6] Juniper Lovato,[1,3] Ellen W. McGinnis,[7] Ryan S. McGinnis,[7] Taylor Ricketts,[4,8] Kathryn Stanton,[6] Peter Sheridan Dodds,[1,9,3,10] and Christopher M. Danforth[1,2,3]

[1] *Vermont Complex Systems Center, MassMutual Center of Excellence for Complex Systems and Data Science, University of Vermont, Burlington, Vermont, USA 05405*
[2] *Department of Mathematics and Statistics, University of Vermont, Burlington, Vermont, USA 05405*
[3] *Computational Story Lab, MassMutual Center of Excellence for Complex Systems and Data Science, University of Vermont, Burlington, Vermont, USA 05405*
[4] *Gund Institute for Environment, University of Vermont, Burlington, Vermont, USA 05405*
[5] *Department of Psychological Science, University of Vermont, Burlington, Vermont, USA 05405*
[6] *Project LEMURS, University of Vermont, Burlington, Vermont, USA 05405*
[7] *Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA 27101*
[8] *Rubenstein School of Environment and Natural Resources, University of Vermont, Burlington, Vermont, USA 05405*
[9] *Department of Computer Science, University of Vermont, Burlington, Vermont, USA 05405*
[10] *Santa Fe Institute, New Mexico, USA 87501*

Wearable devices are rapidly improving our ability to observe health-related processes for extended durations in an unintrusive manner. As part of the Lived Experiences Measured Using Rings Study (LEMURS), we collected heart rate measurements using the Oura ring (Gen3) for over 25,000 sleep periods and self-reported mental health indicators from roughly 600 first-year university students in the United States during the fall semester of 2022. Using clustering techniques, we find that the sleeping heart rate curves can be broadly separated into two categories that are mainly differentiated by how far along the sleep period the lowest heart rate is reached. Sleep periods characterized by a longer time to reach the lowest heart rate are also associated with shorter deep and REM sleep and longer light sleep, but not a difference in total sleep duration. Aggregating sleep periods at the individual level, we find that consistently reaching the lowest heart rate later during sleep is a significant predictor of (1) self-reported impairment due to anxiety or depression, (2) a prior mental health diagnosis, and (3) firsthand experience in traumatic events. This association is more pronounced among females. Our results show that the shape of the sleeping heart rate curve, which is only weakly correlated with descriptive statistics such as the average or the minimum heart rate, is a viable but mostly overlooked metric that can help quantify the relationship between sleep and mental health.

I. INTRODUCTION

Sleep is an important component of well-being, with poor sleep leading to impaired function at the individual and societal [1–5] levels. While some aspects of sleep can be easily monitored, such as the time when an individual goes to bed or wakes up, the unconscious nature of sleep requires external monitoring for proper assessment [6]. Polysomnography has been the gold standard for sleep monitoring, but it requires measurements to be made in controlled conditions in a clinical setting, making it costly and inconvenient [7, 8]. In comparison, the growing availability of consumer-grade wearable devices allows individuals to monitor their sleep for extended periods of time in a non-disruptive and more affordable manner without the recall bias characteristic of self-reports [9]. Feedback from wearable devices is also available soon after data collection, which aids in delivering and assessing the effects of behavioral changes such as interventions.

Sleep has mostly been assessed in terms of sleep duration, sleep efficiency, and the time spent in different sleep stages [10, 11]. Heart rate variability (HRV) is also of interest, particularly in studies on sleep apnea [12–14], and time- and frequency-domain analysis techniques are commonly used [15, 16]. Clustering techniques have been applied on these metrics to find sleep phenotypes, which are sleep patterns shared by a group of individuals that may also share similar characteristics [17–20]. While most use descriptive statistics of these metrics in the clustering algorithm, more recent work used features extracted from the raw sleep-wake time series from more than 100,000 individuals to infer sleep phenotypes related to insomnia [21].

While it has long been known that heart rate generally decreases during sleep [22], the patterns of change in the sleeping heart rate are not well-studied. Unlike other sleep metrics such as sleep stages or heart rate variability, heart rate is more reliably measured by consumer wearable devices [23–27] using photoplethysmography [28], especially when the participant is at rest [29]. A study on heart rate patterns during sleep may give us

* mikaela.fudolig@uvm.edu

valuable insights and these findings may be more consistent across different brands of wearable devices compared to those obtained using other sleep metrics. Specifically, we are interested in extracting sleep phenotypes based on the raw heart rate time series, which gives more information than just the mean or the minimum heart rate.

We would also like to see how these sleeping heart rate patterns relate to mental health, another important component of well-being. Mental health has been shown to have a bidirectional relationship with sleep [5, 30, 31], indicating that treating sleep disorders may also improve mental health [32]. Most studies on mental health and sleep focus on perceived sleep quality [33] and the common sleep metrics described above, such as sleep duration [34–37], heart rate variability [16, 38], or descriptive statistics (mean or minimum) of the heart rate [39–42]. Published studies on the relationship between the shape of the heart rate curve and mental health have focused on its periodicity in the context of daily circadian rhythms [43–46] but not on how heart rate changes during the sleep period itself.

The peak ages for the onset of several mental health conditions, including disorders such as anxiety, depression, and trauma-related conditions, occur before age 25 [47], with around half of the cases manifesting by age 14 [48]. Young adulthood is a critical life stage in detecting and treating mental health conditions and has been the subject of several studies involving mental health [49–52]. An ongoing study at a university in the northeast United States (the Lived Experiences Measured Using Rings Study, or LEMURS) [53–55] monitors the mental and physical well-being of a cohort of college students using surveys and the Oura ring (Gen3), a wearable sleep and activity tracker [23, 24]. The Oura ring provides continuously monitored biometric data, including heart rate, while the surveys regularly collect information about participants' mental health. These self-reported mental health indicators include prior mental health diagnoses, perceived impairment due to anxiety and depression, traumatic events experienced, and stress and anxiety levels. With around half of the study participants having been diagnosed with a mental health condition or having had firsthand experience in two or more types of traumatic events [56], this dataset provides a unique perspective on the relationship between mental health and sleep for a population at risk.

Using data from the LEMURS study, we look at different patterns of change in the heart rate over a sleep period and relate these to the reported mental health indicators of the participants. As we are interested in how heart rate changes across a sleep period, we look at a sleep period not in absolute time (i.e., hours of sleep) but as progressing from the beginning (0% of sleep completed) to the end (100% of sleep completed). Similar to the systematic characterization of sleep phenotypes from sleep-wake measurements using wearable devices [21], we perform clustering algorithms on the heart rate time series to see if these heart rate patterns can be grouped into categories. Associations between these categories and clinically relevant information, such as mental health outcomes were then examined.

II. RESULTS

We processed heart rate (HR) measurements from $m = 20,167$ sleep periods of $N = 599$ participants using piecewise aggregate approximation (PAA) [57–59], which converts the time series from the different sleep periods into equal lengths. After standardizing each time series, we performed k-means clustering, which yields groups of time series with similar shapes. Highly consistent cluster labels were obtained across different centroid initializations or training subsets (each with a size of 10% of the total number of sleep periods) for $k = 2$ clusters but not for higher values of k (Figure 1). When using the full dataset to train the clustering model, 99.96% of the sleep periods had identical cluster assignments after performing k-means 30 times, each with different centroid initializations. Much lower values are obtained for higher values of k (Figures S2 and S3). This separation into two groups was also evident from pairwise correlation maps (Figure S4). We got similar results using constrained dynamic time warping (cDTW) as the measure for k-means, indicating that allowing for slight scaling and translation does not alter our conclusions (Figure S5).

Each cluster was then characterized by the properties of the sleep period and the individual. We compared the clusters using the other sleep period measurements taken by the Oura ring that are not derived from the PAA-processed heart rate time series. We then looked at how the clusters relate to the participant's demographic information, weekly scores in the Perceived Stress Scale (PSS) [60] and the 7-item Generalized Anxiety Disorder Scale (GAD-7) [61], prior mental health diagnoses, prior traumatic events experienced [56], and self-reported effects of anxiety or depression on social or work life. Further, we examined whether the cluster membership of an individual's sleep periods is a significant predictor of an individual's mental health indicators.

A. Cluster characteristics

To understand which aspects of sleep characterize these two clusters, we used a logistic regression model with the cluster label as the response and the sleep period measures recorded by the Oura ring as the predictors. We also compare the distributions of the sleep period measures across the two clusters to examine practical significance.

The two clusters are most differentiated by how far along the sleep period the lowest HR is measured. It is the most relevant predictor in the regression model, resulting in the lowest AIC and residual deviance when

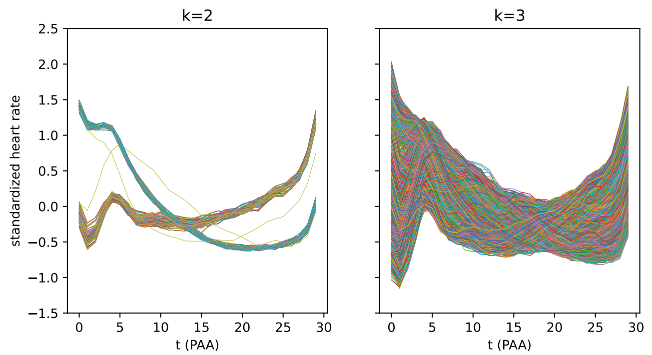

FIG. 1. **Cluster consistency across different training subsets and initializations.** Each curve shows a cluster centroid found for a given number of clusters k using a randomized 10% subset of the data and a randomized centroid initialization to run k-means. Highly consistent cluster centroids are found for $k = 2$, but are not found for $k = 3$ and higher (see Figure S1 for $k > 3$).

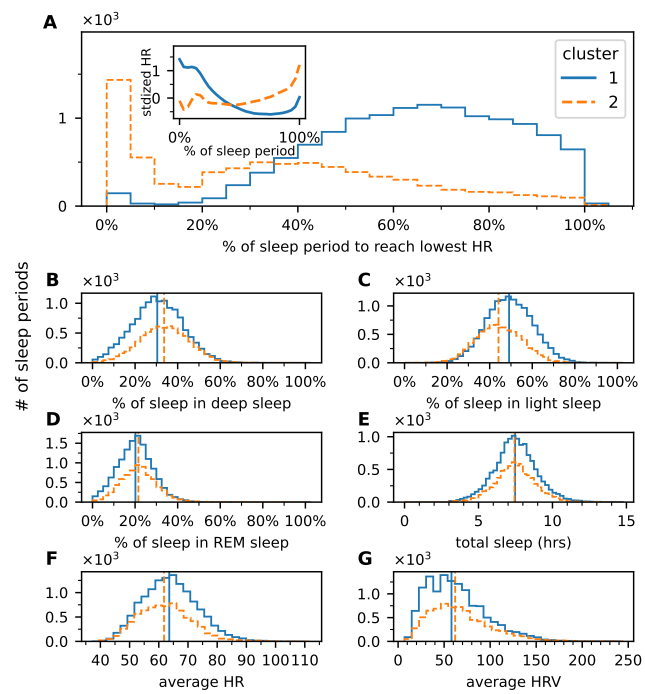

FIG. 2. **Characterizing clusters of sleeping heart rate curves.** The two clusters found are most clearly differentiated by the time it takes to reach the lowest heart rate during the sleep period (A) as compared to other sleep measures taken by the Oura ring. This is consistent with the centroids of the time series in each cluster (inset). The clusters also differ in their sleep stage compositions (B–D) but not in the total sleep duration (E). Smaller but statistically significant differences are also observed for the average heart rate and the average HRV (F–G). Vertical lines indicate the medians for each distribution.

used as a lone predictor, and is the most highly correlated with the cluster label. We can see this clearly in the distributions of this variable for the two clusters, as well as from the shape of the cluster centroids (Figure 2a). The time required to reach the lowest heart rate is only weakly correlated with the average heart rate (Spearman's $\rho = 0.1$ for both the absolute time and the fraction of the sleep period) and the lowest heart rate (Spearman's $\rho = 0.03$ for the absolute time and $\rho = 0.05$ for the fraction of the sleep period). We denote the cluster with the longer median time to reach the lowest heart rate as cluster 1, and the other cluster as cluster 2, corresponding to 64% and 36% of the sleep periods examined, respectively.

The two clusters also differ in terms of sleep stage composition. Cluster 1 is associated with shorter durations of deep (median percentage of the sleep period, 30% vs. 34%) and REM (20% vs. 22%) sleep and longer durations of light sleep (49% vs. 44%). There is no obvious difference in the distributions of the total sleep duration between the two clusters, which is confirmed by the Mann-Whitney U test ($p = 0.18$, Figure 2e). Sleep periods in cluster 1 are also characterized by slightly longer sleep latency, earlier bedtime start and earlier bedtime end, higher average heart rate and lower average HRV. As the average HRV is highly negatively correlated with the average heart rate (Table S3), we do not include it in the regression model. While the average respiratory rate variability was statistically significant in the regression model, we did not find this difference between the two clusters to be practically significant. Differences in the means and medians of the sleep metrics between the two clusters are summarized in Table I.

B. Mental health and sleeping heart rate curves

Now that we have seen which aspects of sleep characterize a cluster, we look at whether the cluster of a given sleep period is related to the individual's demographic information and mental health indicators. As the stress and anxiety scores were collected weekly while the Oura sleep measures was collected nightly, we restrict our data to individuals with at least 3 recorded sleep periods for each week with a survey response. Further, we require at least 10 recorded sleep periods for the duration of the study. This yields $m = 15,073$ sleep periods from $N = 505$ participants.

We use a mixed-effects logistic regression model with the cluster label of a sleep period as the response and the participant ID as a random effect. Adding the week number as either a random or a fixed effect, or excluding it entirely, resulted in similar coefficients and p-values for the other predictors. For this reason, we omit the week number in our final models.

The following predictors were considered for the fixed effects, with the composition among the participants giv-

TABLE I. Mean and median sleep measures for each cluster

	mean		median	
sleep measure / cluster	**1**	**2**	**1**	**2**
lowest HR time offset (% of sleep period)	65	35	66	34
% of sleep period in deep sleep	31	34	30	34
% of sleep period in REM sleep	20	22	20	22
% of sleep period in light sleep	49	45	49	44
total sleep duration (hrs)	7.49	7.46	7.48	7.43
sleep efficiency (%)	87.81	88.54	89	89
sleep latency (mins)	12.91	8.97	9	7
bedtime start (hrs from 12mn)	0.46	0.63	0.38	0.54
bedtime end (hrs from 12mn)	9.01	9.06	8.92	8.98
bedtime duration	8.54	8.44	8.52	8.43
average respiratory rate	15.64	15.5	15.5	15.38
average respiratory rate variation	3.35	3.36	3.25	3.25
average heart rate	63.94	62.22	63.59	61.8
average HRV	63.32	67.18	58	62

en in parentheses (Table S1): gender (68% female, 26% male, 6% other genders), race (88% white, 12% non-white), PSS scores, GAD-7 scores, existence of a prior mental health diagnosis (45% with, 55% without), the number of traumatic event categories experienced first-hand [56] (28% with 0, 32% with 1, 40% with $\geq$ 2), and the existence of perceived effects of diagnosed anxiety or depression on social or work activities (35% with, 65% without). We can interpret the latter as perceived impairment due to anxiety or depression. With 96% of those who reported a prior diagnosis reporting an anxiety or depression diagnosis, the absence of perceived impairment due to anxiety and depression is a good indicator for the vast majority of participants that the diagnosed condition has been controlled.

While the existence of a prior mental health diagnosis and perceived impairment due to anxiety or depression are very highly correlated (Spearman's ρ=0.8) and should not be used in the same model, we can test which of these two predictors produces a better fit. PSS scores and GAD-7 scores are also highly correlated (Spearman's ρ=0.7) so we only use one in any given model to determine the effect of stress or anxiety levels.

PSS scores, GAD-7 scores, and race are not statistically significant at $\alpha = 0.05$, regardless of whether they are used as lone predictors or used in conjunction with others. Having a prior mental health diagnosis, perceived impairment due to anxiety or depression, or firsthand experience in 2 or more types of traumatic events are highly significant ($p \leq 0.001$) as lone predictors of the sleep period cluster. Gender is not statistically significant at $\alpha = 0.05$ when used as a lone predictor or used together with traumatic events experienced, but is statistically significant when used together with perceived impairment or prior mental health diagnoses. Interaction effects between gender and perceived impairment, as well as gender and traumatic events experienced, were not statistically significant when the main effects are also included.

For the final model, we use perceived impairment, first-hand experience in 2 or more types of traumatic events, and gender as the fixed effects. Having a perceived impairment was highly associated ($p = 0.001$) with being in cluster 1, which is characterized by a longer time to reach the lowest HR, as is having firsthand experience in 2 or more types of traumatic events ($p = 0.013$). Being male is also associated with higher odds of being in cluster 1 ($p = 0.028$).

As a post-hoc analysis, we also ran the regression models separately for males, females, and those who do not identify as either gender. While "non-binary" is only one of the gender categories presented in the survey outside of "male" and "female", for brevity, we will refer to all these other categories as "non-binary". Perceived impairment is associated with increased odds of being in cluster 1 in sleep periods among females ($p < 0.001$), while the number of types of traumatic events experienced first-hand (whether < 2 or ≥ 2) is associated with increased odds of being in cluster 1 in sleep periods among females ($p = 0.008$) and non-binary individuals ($p = 0.013$). Both variables are not statistically significant for males ($p = 0.622$ and $p = 0.309$ for perceived impairment and existence of traumatic events, respectively).

In addition, if we restrict the data to sleep periods of those with or without perceived impairment, we find that the odds of being in cluster 1 among those without impairment are higher for males than for females. No gender effect is seen among those with perceived impairment. On the other hand, gender is only a significant predictor for those with two or more types of traumatic events, with non-binary individuals having sleep periods of higher odds of being in cluster 1, but not for those who experienced fewer than two. Detailed regression results are given in Tables S5–S7.

TABLE II. Impairment, trauma, and cluster consistency by gender. `frac_1` is the fraction of sleep periods in cluster 1.

	Imp	indiv		nights		frac_1		**Trau**	indiv		nights		frac_1	
gender		size	pct	size	pct	mean	median		size	pct	size	pct	mean	median
f (68%)	**0**	212	61.6%	6584	62.9%	0.59	0.62	$<$ **2**	204	59.3%	6443	61.5%	0.60	0.62
	1	132	38.4%	3885	37.1%	0.68	0.72	$\geq$ **2**	140	40.7%	4026	38.5%	0.66	0.69
m (26%)	**0**	106	82.2%	3027	84.1%	0.66	0.68	$<$ **2**	83	64.3%	2359	65.5%	0.65	0.67
	1	23	17.8%	574	16.0%	0.69	0.72	$\geq$ **2**	46	35.7%	1242	34.5%	0.70	0.71
nb (6%)	**0**	11	34.4%	388	38.7%	0.57	0.66	$<$ **2**	18	56.3%	584	58.2%	0.60	0.65
	1	21	65.6%	615	61.3%	0.74	0.76	$\geq$ **2**	14	43.8%	419	41.8%	0.79	0.88

C. Predicting an individual's mental health indicator from sleep period cluster data

With the main predictors not related to any of the weekly measures, we can aggregate our data at the individual level and predict mental health indicators from how often their sleep periods belong in a given cluster. Using logistic regression, we find that having a higher fraction of sleep periods in cluster 1 is associated with higher odds of having perceived impairment due to anxiety or depression ($p < 0.001$), a prior mental health diagnosis ($p = 0.003$), or firsthand experience in 2 or more traumatic event categories ($p = 0.001$). It is a statistically significant predictor of being female (vs. non-female, $p = 0.045$). Still, this fraction is not significant in differentiating males and non-males ($p = 0.115$) or non-binary individuals and those who are either male or female ($p = 0.31$).

As gender is common demographic information to ask in health-related apps, we check whether this aids in predicting mental health indicators when used together with the fraction of sleep periods in cluster 1. This fraction and gender are statistically significant in predicting perceived impairment and prior mental health diagnoses ($p < 0.01$). However, in predicting firsthand experience in 2 or more traumatic events, gender is not statistically significant.

We also perform a post-hoc analysis to dig deeper into the effect of gender. The fraction of sleep periods belonging to cluster 1 is a statistically significant predictor for perceived impairment due to anxiety or depression only for females; for traumatic events, it is statistically significant only for females and non-binary individuals. In these cases, more sleep periods in cluster 1 results in higher odds of having an impairment or firsthand experience in two or more types of traumatic events. This mirrors our earlier post-hoc analysis in predicting the sleep period cluster based on the individual's characteristics. One can get an intuition of the regression results from the composition of our data as given in Table II and Figure S7. Detailed regression results are given in Tables S8–S9.

III. DISCUSSION

Using measurements taken from first-year university students using the Oura sleep tracker, we study patterns in how the heart rate changes over a sleep period and how this relates to the individual's mental health indicators. We find two broad categories of sleeping heart rate curves mainly differentiated by when the lowest heart rate is attained. Sleep periods where the lowest heart rate is reached later are also characterized by shorter deep and REM sleep and longer light sleep; no significant difference between the two categories was found in total sleep duration. Sleep periods belonging to individuals who self-report impairment due to anxiety or depression, or those who have experienced firsthand two or more named categories of traumatic events [56], are more likely to attain the lowest heart rate later in sleep. In contrast, weekly stress and anxiety levels, as measured through the PSS or GAD-7 scores, are not associated with sleeping heart rate changes. [54].

We find differences across genders in how perceived impairment and traumatic events are associated with the sleep period cluster. Specifically, while the sleep periods of females differ between those with perceived impairment or firsthand experience in two or more types of traumatic events and those without, this does not apply to the sleep periods of males. Sleep periods of non-binary individuals differ between those with firsthand experience of two or more traumatic events and those with fewer. Still, there is no significant difference between those with perceived impairment and those without.

As the sleeping heart rate curve pattern only depends on the properties of the individual in our data, we also look at reversing the prediction task, i.e., whether the sleeping heart rate curve can be used to predict the mental health state of the participant. The fraction of sleep periods of a participant in a given pattern is a significant predictor for the existence of perceived impairment due to anxiety or depression or a prior mental health diagnosis, as well as the number of traumatic event categories experienced firsthand. These further support the relationship we found earlier between sleeping heart rate patterns and mental health.

We also observe a gender effect consistent with what we found in the models predicting sleep period clusters.

The fraction of sleep periods in a given pattern is significant in predicting perceived impairment due to anxiety or depression only for females. It is also significant only among females and non-binary individuals for predicting whether the number of types of traumatic events experienced firsthand is two or higher.

While studies tend to aggregate heart measurements into a single statistic, in particular the average and the minimum, these are only weakly correlated with the pattern of change in the sleeping heart rate. Our results show that not only do we find different patterns in how the heart rate changes during sleep, but also that these patterns are related to mental health, indicating that the shape of the heart rate curve is a viable but underexplored sleep metric.

To our knowledge, the only other study on the shapes of sleeping heart rate curves was performed by Oura [62]. They found four shapes, contrasting with the two broad categories for which we found strong support. While the methodology and analysis in this study were not released, we were able to reproduce these four shapes by setting $k = 4$ in the k-means algorithm and taking the cluster centroids (Figure S3). However, for our dataset, setting $k = 4$ results in less consistent cluster labels with different training subsets or initializations, unlike the case of $k = 2$. Our analysis points to a spectrum of shapes, including the four shapes that Oura found, that can be broadly categorized into two groups. It is worth noting that Oura's analysis reflects several orders of magnitude more individuals, with far broader demographic variation.

The existence of two robust clusters suggests that there may be sleep phenotypes in relation to the sleeping heart rate. Sleep phenotypes have recently been inferred using clustering techniques on large-scale accelerometer sleep/wake time series data [21], particularly focusing on insomnia. Sleep patterns have also been identified for other mental health diagnoses [63–65]. With the majority of the mental health diagnoses in our sample being anxiety or depression, our findings also point to links between these disorders and how heart rate changes during sleep.

The relationship between mental health, particularly anxiety, depression, and trauma, and the consistency in sleeping heart rate curve patterns is of interest, particularly since the heart rate curves are most correlated to the time to reach the lowest heart rate in sleep, a sleep measure that is not commonly used. One possible mechanism could be by directly affecting the nervous system. For example, some mental health disorders, such as anxiety, have been known to affect neurotransmitters [66, 67] and brain regions [68] that regulate sleep. Anxiety disorders are also associated with amygdala hyperactivity [69, 70], which in turn is linked to heart rate variability [71]. However, mental health disorders are also associated with habits that affect sleep, such as food intake [72, 73]. Adding other wearable devices, such as glucose monitors for monitoring food intake, to sleep studies [74, 75] is a promising direction to test these hypotheses.

We also observe interesting gender differences in how the sleeping heart rate curve relates to mental health. While females have different sleeping heart rate curve shapes depending on whether they have anxiety, depression, or past trauma, this is not observed among males. Previous studies [76] have mainly shown that females sleep for longer and have longer deep sleep but also report poorer sleep quality [77–79]. Regardless, there are conflicting results [80], possibly due to different study designs. With regard to the sleeping heart rate curve, we find that males, whether or not they had mental health disorders or trauma, have higher odds of having sleeping heart rate curves that correspond to a longer time to reach the lowest heart rate. On the other hand, females who do not have mental health diagnoses or firsthand experience in 2 or more types of traumatic events are associated with sleep periods where the lowest heart rate is attained earlier at night. Given that in this same dataset, females have longer sleep duration, albeit similar deep sleep, than males (Table S10), the sleeping heart rate curve provides different information from just the total sleep duration or deep sleep alone. We also note that our sample is highly homogeneous in age, which is not the case in most research on gender differences in sleep.

While several studies show that stress and anxiety levels affect sleep [40, 81–83], our data shows that weekly PSS and GAD-7 scores are not statistically significant in predicting the sleeping heart rate curve pattern. This may be due to a difference in temporal resolution: sleep periods are monitored nightly, while the stress and anxiety levels are only obtained from a weekly self-report [54]. We are implementing more frequent surveys in future studies to ascertain whether daily fluctuations in stress or anxiety relate to sleeping heart rate patterns.

We are interested in studying university students as they are a population at risk for which mental health interventions are highly relevant and more easily implementable [53–55]. Our sample has a high proportion (45%) of students with a prior mental health diagnosis, 96% of whom reporting an anxiety or depression diagnosis. This is in agreement with the estimated prevalence of anxiety among university students [84].

The selection of this narrow demographic is advantageous in several ways, particularly in preventing study dropouts and administering interventions. However, this also limits the generalizability of our results to other age groups. Our sample was also relatively racially homogeneous, which may explain the insignificance of race in our results. While sleep disorders are not uncommon among university students [85, 86], certain sleep orders, such as sleep apnea, are more common in older populations [87]. Further, stressors also change over different life stages [88]. Thus, we expect that certain sleep patterns do not manifest in the age group studied. We hope to address these limitations by expanding our recruitment to include a more diverse set of individuals in future work.

IV. METHODS

A. Data

Sleep data was obtained from 603 participants in Project LEMURS, who, at the time of the study, were first-year students in a university in the United States [53–55]. These participants were asked to wear the Oura ring (Gen3) during sleep for a period of 8 weeks (Oct–Dec 2022). We look at the heart rate time series measured by the ring at 5-minute intervals from the start of the detected sleep period. Naps are excluded by considering only sleep periods marked by Oura as "long sleep." While there is mostly only one such sleep period per participant per night, we exclude 368 sleep periods considered as "long sleep" for the same participant and for the same night. This brings the number of time series to 25,800.

To make the heart rate time series comparable for sleep periods of different durations, we use piecewise aggregate approximation (PAA) [58, 59] and divide a time series into $n = 30$ equally-sized segments. For each segment, we take the mean of the heart rate measurements in that segment as its representative value. This yields a PAA-processed time series of length $n = 30$ regardless of the length of the raw time series. The number of segments was based heuristically on the distribution of the length of the raw heart rate time series across all sleep periods (Figure S6): with a median of 102 data points per sleep period, $n = 30$ roughly corresponds to 15 minutes of sleep per segment. Using $n = 30$ also allows us to include even the shortest sleep period (39 data points, or roughly 3.25 hours of sleep).

As movement during sleep results in poorer heart rate estimates [29], the Oura ring does not provide HR measurements when there is significant movement. This creates missing values in the raw time series. When using PAA, we take the means for each segment, disregarding any missing values in that segment. If all the entries in a given segment are missing, then the corresponding PAA mean of that segment is also considered missing. PAA computations were made with a modified version of the Python package `pyts` [89] that disregards missing data when computing the mean. Sleep periods where there is at least one segment with a missing PAA value are disregarded. This reduces the number of sleep periods examined to $m = 20,167$ belonging to $N = 599$ individuals. The 5,629 disregarded sleep periods, excluding four sleep periods with no heart rate data, span a similar range of sleep period lengths (Figure S6) and are from 587 individuals, indicating that disregarding time series with missing PAA values does not sufficiently discriminate against a few users or sleep period durations.

Participants in the LEMURS study completed baseline surveys at the start of the study (Table S1). The survey asks whether the participant has had a prior diagnosis of the following common mental health conditions as part of a standardized health screening questionnaire for a range of physical and mental health conditions [90]: anxiety, depression, attention-deficit/hyperactivity disorder (ADHD), alcoholism, psychosis, delusions, anorexia or bulimia, post-traumatic stress disorder (PTSD), obsessive-compulsive disorder (OCD), bipolar disorder (BPD), panic attacks, and emotional disorder. If the participant reports an anxiety or depression diagnosis, they are asked whether their condition affects their social or work activities in any way. It also asks participants to report traumatic events experienced named in the Life Events Checklist [56]. We specifically look at the number of categories of named traumatic events experienced firsthand by the participant (0, 1, or ≥ 2).

In addition to these questions, the survey also asks for the gender and race of the participant. The students all entered the university in the same academic year and were required to be within 18-24 years of age to be part of the study. 96% of the students reported birth years between 2002 and 2004. The participants answered the PSS and GAD-7 questionnaires at the end of each week.

B. Clustering the time series

We perform k-means clustering [91] on the $m = 20,167$ heart rate time series processed with PAA and then standardized, each of length $n = 30$. The k-means algorithm is a popular method for time series clustering [92, 93] due to its performance and simplicity. The Python package `tslearn` [94] was used for implementation. The Euclidean distance is a good choice to measure distances, as we are interested in the general shape of the heart rate curve and when certain points, such as the lowest heart rate, are achieved during the sleep period.

To determine the optimum number of clusters k, we use consistency metrics. First, we use 30 randomly selected subsets of the data (10% of the total size) to train the model. For each training subset, we perform k-means with 30 different centroid initializations and assign the rest of the data points to a given cluster based on the trained model. This results in 900 different k-means runs for each value of k. If the clustering is robust, the time series must be clustered similarly for different runs, resulting in the convergence of the cluster centroids across the different randomizations.

Because k-means assigns cluster labels arbitrarily, we have to ensure that the same cluster labels refer to similar cluster centroids for each run. Once we establish that the cluster centroids are consistent across different randomizations, we standardize the cluster centroids obtained by training the k-means model on the entire dataset for 30 different initializations. We perform another round of k-means on these centroids, resulting in what we call "metaclusters." The cluster centroids assigned to the same metacluster will be assigned the same cluster label, and this cluster label will be transmitted back to the raw heart rate time series. Thus, each raw heart rate time series will have 30 cluster labels, and the mode of these will be the final cluster label assignment for the time series. We also record how often a time series is assigned to the same cluster label.

We also used constrained dynamic time warping (cDTW) using Sakoe-Chiba bands [95, 96] of widths $w = 3$ and $w = 6$ (note that our time series is of length $n = 30$) to explore the effects of slight scaling or translation. As the cluster centroids obtained are similar for both cDTW and Euclidean distance (Figure S5), we focus on the results obtained using the Euclidean distance.

C. Understanding the clusters

Aside from the heart rate time series, the Oura ring also collects other sleep measures, including the sleep period start and end times, the estimated durations of each sleep stage (light, REM, and deep sleep) as well as the time spent awake during the sleep period, the average and lowest heart rates, the time when the lowest heart rate is obtained, sleep latency, average respiratory rate, average respiratory rate variation, and heart rate variability. We use generalized linear models to gain insight into the sleep measures that differentiate the clusters. Because the clusters were obtained solely from the shapes of the heart rate curves, we do not add random effects. Regression models were performed with `pymer4` [97] and confirmed using the `glm` model in R. After performing regression models using each measure separately, we combine the sleep measures in a single model and use `stepAIC` of the `MASS` library to perform stepwise regression. We then check for practical significance by examining the distributions and descriptive statistics of each sleep measure for the clusters found.

D. Relating clusters to mental health

For each sleep period, we take the cluster label as the outcome variable and the survey responses of the individual as the predictors. Specifically, we look at the individual's baseline survey responses on prior mental health diagnoses, traumatic events experienced, gender, and race, as well as the individual's weekly responses resulting in their PSS and GAD-7 scores. For those who reported anxiety and depression diagnoses, we also have information on whether the participants think their condition affects their social or work activities in any way. This can be interpreted as perceived impairment due to anxiety or depression.

As several periods correspond to the same individual and the same week, we use a linear mixed-effects model to account for these groupings. Since the PSS and GAD-7 scores are obtained weekly while the sleep data is obtained nightly, we only include weeks where a given participant answered the weekly survey and had at least three (3) recorded sleep periods. We also require that this restriction yields at least ten (10) sleep periods for a given participant. These ensure that the weight of a given sleep period to the weekly PSS score is not artificially high and that enough data points are available for the fraction of sleep periods in a given cluster for each individual to be a reasonable measure. This results in a dataset with 15,073 sleep periods from $N = 505$ participants. Linear mixed-effects models, with the participant ID and the week number treated as random effects, were implemented using the Python package `pymer4` [97], which interfaces with the `glmer` [98] and `lmerTest` [99] packages in R. We note that the participant IDs were treated as categorical variables while the week numbers were treated as integers ranging from 0 to 7.

As the last part of our analysis, we predict mental health indicators at the individual level, in particular, perceived impairment due to anxiety or depression, the presence of a prior mental health diagnosis, and the number of traumatic event categories experienced firsthand from the fraction of sleep periods in a given cluster. We use logistic regression implemented using the Python package `pymer4`.

ACKNOWLEDGMENTS

This study was funded by a grant from MassMutual. The manuscript benefited from helpful discussions with Alec Beauregard and Marco Bonete.

[1] Nicola Magnavita and Sergio Garbarino, "Sleep, Health and Wellness at Work: A Scoping Review," International Journal of Environmental Research and Public Health **14**, 1347 (2017).

[2] Marco Hafner, Martin Stepanek, Jirka Taylor, Wendy M. Troxel, and Christian van Stolk, "Why Sleep Matters—The Economic Costs of Insufficient Sleep," Rand Health Quarterly **6**, 11 (2017).

[3] David Hillman, Scott Mitchell, Jared Streatfeild, Chloe Burns, Dorothy Bruck, and Lynne Pezzullo, "The economic cost of inadequate sleep," Sleep **41**, zsy083 (2018).

[4] Jean-Philippe Chaput, Julie Carrier, Célyne Bastien, Geneviève Gariépy, and Ian Janssen, "Economic burden of insufficient sleep duration in Canadian adults," Sleep Health **8**, 298–302 (2022).

[5] Laura Palagini, Elisabeth Hertenstein, Dieter Riemann, and Christoph Nissen, "Sleep, insomnia and mental health," Journal of Sleep Research **31**, e13628 (2022).

[6] Ruth Ravichandran, Sang-Wha Sien, Shwetak N. Patel, Julie A. Kientz, and Laura R. Pina, "Making Sense of Sleep Sensors: How Sleep Sensing Technologies Support and Undermine Sleep Health," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Sys-*

tems, CHI '17 (Association for Computing Machinery, New York, NY, USA, 2017) pp. 6864–6875.

[7] Jessica M. Kelly, Robert E. Strecker, and Matt T. Bianchi, "Recent Developments in Home Sleep-Monitoring Devices," International Scholarly Research Notices **2012**, e768794 (2012).

[8] Elise Guillodo, Christophe Lemey, Mathieu Simonnet, Michel Walter, Enrique Baca-García, Vincent Masetti, Sorin Moga, Mark Larsen, Hugopsy Network, Juliette Ropars, and Sofian Berrouiguet, "Clinical Applications of Mobile Health Wearable–Based Sleep Monitoring: Systematic Review," JMIR mHealth and uHealth **8**, e10733 (2020).

[9] Diane S. Lauderdale, Kristen L. Knutson, Lijing L. Yan, Kiang Liu, and Paul J. Rathouz, "Sleep duration: how well do self-reports reflect objective measures? The CARDIA Sleep Study," Epidemiology (Cambridge, Mass.) **19**, 838–845 (2008).

[10] Deepak Shrivastava, Syung Jung, Mohsen Saadat, Roopa Sirohi, and Keri Crewson, "How to interpret the results of a sleep study," Journal of Community Hospital Internal Medicine Perspectives **4**, 10.3402/jchimp.v4.24983 (2014).

[11] Katherine A. Kaplan, Jason Hirshman, Beatriz Hernandez, Marcia L. Stefanick, Andrew R. Hoffman, Susan Redline, Sonia Ancoli-Israel, Katie Stone, Leah Friedman, and Jamie M. Zeitzer, "When a gold standard isn't so golden: Lack of prediction of subjective sleep quality from sleep polysomnography," Biological Psychology **123**, 37–46 (2017).

[12] Christian Guilleminault, Roger Winkle, Stuart Connolly, Kenneth Melvin, and Ara Tilkian, "Cyclical variation of the heart rate in sleep apnoea syndrome: mechanisms, and usefulness of 24 h electrocardiography as a screenting technique," The Lancet **323**, 126–131 (1984).

[13] Vanessa Cristina Cunha Sequeira, Pamela Martin Bandeira, and João Carlos Moreno Azevedo, "Heart rate variability in adults with obstructive sleep apnea: a systematic review," Sleep Science **12**, 214–221 (2019).

[14] Seren Ucak, Hasthi U. Dissanayake, Kate Sutherland, Philip de Chazal, and Peter A. Cistulli, "Heart rate variability and obstructive sleep apnea: Current perspectives and novel technologies," Journal of Sleep Research **30**, e13274 (2021).

[15] Thomas Penzel, Jan W. Kantelhardt, Ronny P. Bartsch, Maik Riedl, Jan F. Kraemer, Niels Wessel, Carmen Garcia, Martin Glos, Ingo Fietze, and Christoph Schöbel, "Modulations of Heart Rate, ECG, and Cardio-Respiratory Coupling Observed in Polysomnography," Frontiers in Physiology **7** (2016).

[16] Arron T. L. Correia, Gosia Lipinska, H. G. Laurie Rauch, Philippa E. Forshaw, Laura C. Roden, and Dale E. Rae, "Associations between sleep-related heart rate variability and both sleep and symptoms of depression and anxiety: A systematic review," Sleep Medicine **101**, 106–117 (2023).

[17] Christopher B. Miller, Delwyn J. Bartlett, Anna E. Mullins, Kirsty L. Dodds, Christopher J. Gordon, Simon D. Kyle, Jong Won Kim, Angela L. D'Rozario, Rico S.C. Lee, Maria Comas, Nathaniel S. Marshall, Brendon J. Yee, Colin A. Espie, and Ronald R. Grunstein, "Clusters of Insomnia Disorder: An Exploratory Cluster Analysis of Objective Sleep Parameters Reveals Differences in Neurocognitive Functioning, Quantitative EEG, and Heart Rate Variability," Sleep **39**, 1993–2004 (2016).

[18] Vivian W. Chiu, Melissa Ree, Aleksandar Janca, Rajan Iyyalol, Milan Dragovic, and Flavie Waters, "Sleep profiles and CBT-I response in schizophrenia and related psychoses," Psychiatry Research **268**, 279–287 (2018).

[19] Andrey V. Zinchuk, Sangchoon Jeon, Brian B. Koo, Xiting Yan, Dawn M. Bravata, Li Qin, Bernardo J. Selim, Kingman P. Strohl, Nancy S. Redeker, John Concato, and Henry K. Yaggi, "Polysomnographic phenotypes and their cardiovascular implications in obstructive sleep apnoea," Thorax **73**, 472–480 (2018).

[20] Adriana Kramer Fiala Machado, Andrea Wendt, Ana Maria Baptista Menezes, Helen Gonçalves, and Fernando C. Wehrmeister, "Sleep clusters and modifiable risk behaviors for noncommunicable diseases in young adults: Data from a birth cohort in Brazil," Sleep Health **9**, 346–353 (2023).

[21] Machiko Katori, Shoi Shi, Koji L. Ode, Yasuhiro Tomita, and Hiroki R. Ueda, "The 103,200-arm acceleration dataset in the UK Biobank revealed a landscape of human sleep phenotypes," Proceedings of the National Academy of Sciences **119**, e2116729119 (2022).

[22] Frederick Snyder, J. Allan Hobson, Donald F. Morrison, and Frederick Goldfrank, "Changes in respiration, heart rate, and systolic blood pressure in human sleep," Journal of Applied Physiology **19**, 417–422 (1964).

[23] Massimiliano de Zambotti, Leonardo Rosas, Ian M. Colrain, and Fiona C. Baker, "The Sleep of the Ring: Comparison of the ŌURA Sleep Tracker Against Polysomnography," Behavioral Sleep Medicine **17**, 124–136 (2019).

[24] Marco Altini and Hannu Kinnunen, "The Promise of Sleep: A Multi-Sensor Approach for Accurate Sleep Stage Detection Using the Oura Ring," Sensors **21**, 4302 (2021).

[25] Clint R. Bellenger, Dean J. Miller, Shona L. Halson, Gregory D. Roach, and Charli Sargent, "Wrist-Based Photoplethysmography Assessment of Heart Rate and Heart Rate Variability: Validation of WHOOP," Sensors **21**, 3571 (2021).

[26] Syed Anas Imtiaz, "A systematic review of sensing technologies for wearable sleep staging," Sensors **21**, 1562 (2021).

[27] Rui Cao, Iman Azimi, Fatemeh Sarhaddi, Hannakaisa Niela-Vilen, Anna Axelin, Pasi Liljeberg, and Amir M. Rahmani, "Accuracy assessment of Oura Ring nocturnal heart rate and heart rate variability in comparison with electrocardiography in time and frequency domains: comprehensive analysis," Journal of Medical Internet Research **24**, e27487 (2022).

[28] Denisse Castaneda, Aibhlin Esparza, Mohammad Ghamari, Cinna Soltanpur, and Homer Nazeran, "A review on wearable photoplethysmography sensors and their potential future applications in health care," International journal of biosensors & bioelectronics **4**, 195–202 (2018).

[29] Brinnae Bent, Benjamin A. Goldstein, Warren A. Kibbe, and Jessilyn P. Dunn, "Investigating sources of inaccuracy in wearable optical heart rate sensors," npj Digital Medicine **3**, 1–9 (2020).

[30] Dan Robotham, "Sleep as a public health concern: insomnia and mental health," Journal of Public Mental Health **10**, 234–237 (2011).

[31] Xiaohui Sun, Bin Liu, Sitong Liu, David J. H. Wu, Jianming Wang, Yi Qian, Ye Ding, and Yingying Mao, "Sleep disturbance and psychiatric disorders: a bidirectional Mendelian randomisation study," Epidemiology and Psychiatric Sciences **31** (2022), 10.1017/S2045796021000810.

[32] Daniel Freeman, Bryony Sheaves, Felicity Waite, Allison G Harvey, and Paul J Harrison, "Sleep disturbance and psychiatric disorders," The Lancet Psychiatry **7**, 628–637 (2020).

[33] Holly J. Ramsawh, Murray B. Stein, Shay-Lee Belik, Frank Jacobi, and Jitender Sareen, "Relationship of anxiety disorders, sleep quality, and functional impairment in a community sample," Journal of Psychiatric Research **43**, 926–933 (2009).

[34] Roger R. Rosa, Michael H. Bonnet, and Milton Kramer, "The relationship of sleep and anxiety in anxious subjects," Biological Psychology **16**, 119–126 (1983).

[35] Kristi H. Fuller, William F. Waters, Paul G. Binks, and Tai Anderson, "Generalized Anxiety and Sleep Architecture: A Polysomnographic Investigation," Sleep **20**, 370–376 (1997).

[36] Josine G. Van Mill, Witte J. G. Hoogendijk, Nicole Vogelzangs, Richard Van Dyck, and Brenda W. J. H. Penninx, "Insomnia and Sleep Duration in a Large Cohort of Patients With Major Depressive Disorder and Anxiety Disorders," The Journal of Clinical Psychiatry **71**, 239–246 (2010).

[37] Chiara Baglioni, Svetoslava Nanovska, Wolfram Regen, Kai Spiegelhalder, Bernd Feige, Christoph Nissen, Charles F. Reynolds, and Dieter Riemann, "Sleep and mental disorders: A meta-analysis of polysomnographic research," Psychological Bulletin **142**, 969–990 (2016).

[38] Gaetano Valenza, Luca Citi, Antonio Lanata, Claudio Gentili, Riccardo Barbieri, and Enzo Pasquale Scilingo, "Applications of Heartbeat Complexity Analysis to Depression and Bipolar Disorder," in *Complexity and Nonlinearity in Cardiovascular Signals*, edited by Riccardo Barbieri, Enzo Pasquale Scilingo, and Gaetano Valenza (Springer International Publishing, Cham, 2017) pp. 345–374.

[39] Heli Koskimäki, Hannu Kinnunen, Salla Rönkä, and Benjamin Smarr, "Following the heart: what does variation of resting heart rate tell about us as individuals and as a population," in *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, UbiComp/ISWC '19 Adjunct (Association for Computing Machinery, New York, NY, USA, 2019) pp. 1178–1181.

[40] Yasmine Azza, Marcus Grueschow, Walter Karlen, Erich Seifritz, and Birgit Kleim, "How stress affects sleep and mental health: nocturnal heart rate increases during prolonged stress and interacts with childhood trauma exposure to predict anxiety," Sleep **43**, zsz310 (2020).

[41] Steven H. Woodward, Andrea L. Jamison, Sasha Gala, Catherine Lawlor, Diana Villasenor, Gisselle Tamayo, and Melissa Puckett, "Heart rate during sleep in PTSD patients: Moderation by contact with a service dog," Biological Psychology **180**, 108586 (2023).

[42] S. Siddi, R. Bailon, I. Giné-Vázquez, F. Matcham, F. Lamers, S. Kontaxis, E. Laporta, E. Garcia, F. Lombardini, P. Annas, M. Hotopf, B. W. J. H. Penninx, A. Ivan, K. M. White, S. Difrancesco, P. Locatelli, J. Aguiló, M. T. Peñarrubia-Maria, V. A. Narayan, A. Folarin, D. Leightley, N. Cummins, S. Vairavan, Y. Ranjan, A. Rintala, G. De Girolamo, S. K. Simblett, T. Wykes, PAB members, I. Myin-Germeys, R. Dobson, and J. M. Haro, "The usability of daytime and night-time heart rate dynamics as digital biomarkers of depression severity," Psychological Medicine **53**, 3249–3260 (2023).

[43] Jacques Taillard, Patrick Lemoine, Pierre Boule, Michel Drogue, and Jacques Mouret, "Sleep and Heart Rate Circadian Rhythm in Depression: The Necessity to Separate," Chronobiology international **10**, 63–72 (1993).

[44] Hans G. Stampfer, "The Relationship between Psychiatric Illness and the Circadian Pattern of Heart Rate," Australian & New Zealand Journal of Psychiatry **32**, 187–198 (1998).

[45] Ian B. Hickie, Sharon L. Naismith, Rébecca Robillard, Elizabeth M. Scott, and Daniel F. Hermens, "Manipulating the sleep-wake cycle and circadian rhythms to improve clinical management of major depression," BMC Medicine **11**, 79 (2013).

[46] Erika Lutin, Carmen Schiweck, Jan Cornelis, Walter De Raedt, Andreas Reif, Elske Vrieze, Stephan Claes, and Chris Van Hoof, "The cumulative effect of chronic stress and depressive symptoms affects heart rate in a working population," Frontiers in Psychiatry **13** (2022).

[47] Marco Solmi, Joaquim Radua, Miriam Olivola, Enrico Croce, Livia Soardo, Gonzalo Salazar de Pablo, Jae Il Shin, James B. Kirkbride, Peter Jones, Jae Han Kim, Jong Yeob Kim, Andrè F. Carvalho, Mary V. Seeman, Christoph U. Correll, and Paolo Fusar-Poli, "Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies," Molecular Psychiatry **27**, 281–295 (2022).

[48] Ronald C. Kessler, Patricia Berglund, Olga Demler, Robert Jin, Kathleen R. Merikangas, and Ellen E. Walters, "Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication," Archives of General Psychiatry **62**, 593–602 (2005).

[49] Edward M. Adlaf, Louis Gliksman, Andrée Demers, and Brenda Newton-Taylor, "The Prevalence of Elevated Psychological Distress Among Canadian Undergraduates: Findings from the 1998 Canadian Campus Survey," Journal of American College Health **50**, 67–72 (2001).

[50] Stephen P. Becker, Matthew A. Jarrett, Aaron M. Luebbe, Annie A. Garner, G. Leonard Burns, and Michael J. Kofler, "Sleep in a large, multi-university sample of college students: sleep problem prevalence, sex differences, and mental health correlates," Sleep Health **4**, 174–181 (2018).

[51] Kiera Louise Adams, Kate E. Saunders, Charles Donald George Keown-Stoneman, and Anne C. Duffy, "Mental health trajectories in undergraduate students over the first year of university: a longitudinal cohort study," BMJ Open **11**, e047393 (2021).

[52] William E. Copeland, Ellen McGinnis, Yang Bai, Zoe Adams, Hilary Nardone, Vinay Devadanam, Jeffrey Rettew, and Jim J. Hudziak, "Impact of COVID-19 Pandemic on College Student Mental Health and Wellness," Journal of the American Academy of Child & Adolescent Psychiatry **60**, 134–141.e2 (2021).

[53] Matthew Price, Johanna E. Hidalgo, Yoshi M. Bird, Laura S. P. Bloomfield, Casey Buck, Janine Cerutti, Peter Sheridan Dodds, Mikaela Irene Fudolig, Rachel

Gehman, Marc Hickok, Julia Kim, Jordan Llorin, Juniper Lovato, Ellen W. McGinnis, Ryan S. McGinnis, Richard Norton, Vanessa Ramirez, Kathryn Stanton, Taylor H. Ricketts, and Christopher M. Danforth, "A large clinical trial to improve well-being during the transition to college using wearables: The Lived Experiences Measured Using Rings Study," Contemporary Clinical Trials **133**, 107338 (2023).

[54] Laura Bloomfield, Mikaela I. Fudolig, Peter Sheridan Dodds, Julia Kim, Jordan Llorin, Juniper L. Lovato, Ellen McGinnis, Ryan S. McGinnis, Matthew Price, Taylor H. Ricketts, Kathryn Stanton, and Christopher M. Danforth, "Detecting stress in college freshmen from wearable sleep data," (2023), PsyArXiv:eu896.

[55] Laura Bloomfield, Mikaela Fudolig, Peter Dodds, Julia Kim, Jordan Llorin, Juniper Lovato, Ellen McGinnis, Ryan McGinnis, Matthew Price, Taylor Ricketts, Kathryn Stanton, and Christopher Danforth, "Events and Behaviors Associated with Symptoms of Generalized Anxiety Disorder in First-year College Students," (2023), PsyArXiv:278ey.

[56] Matt J. Gray, Brett T. Litz, Julie L. Hsu, and Thomas W. Lombardo, "Psychometric Properties of the Life Events Checklist," Assessment **11**, 330–341 (2004).

[57] Byoung-Kee Yi and Christos Faloutsos, "Fast Time Sequence Indexing for Arbitrary Lp Norms," in *Proceedings of the 26th International Conference on Very Large Data Bases*, VLDB '00 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2000) pp. 385–394.

[58] Eamonn Keogh, Kaushik Chakrabarti, Michael Pazzani, and Sharad Mehrotra, "Locally adaptive dimensionality reduction for indexing large time series databases," in *Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data*, SIGMOD '01 (Association for Computing Machinery, New York, NY, USA, 2001) pp. 151–162.

[59] Jessica Lin, Eamonn Keogh, Li Wei, and Stefano Lonardi, "Experiencing SAX: a novel symbolic representation of time series," Data Mining and Knowledge Discovery **15**, 107–144 (2007).

[60] Sheldon Cohen, Tom Kamarck, and Robin Mermelstein, "A Global Measure of Perceived Stress," Journal of Health and Social Behavior **24**, 385–396 (1983).

[61] Robert L. Spitzer, Kurt Kroenke, Janet B. W. Williams, and Bernd Löwe, "A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7," Archives of Internal Medicine **166**, 1092–1097 (2006).

[62] Oura Team, "Sleeping Heart Rate: Look for These 3 Patterns," (2023).

[63] Silvia Miano, Pasquale Parisi, and Maria Pia Villa, "The sleep phenotypes of attention deficit hyperactivity disorder: The role of arousal during sleep and implications for treatment," Medical Hypotheses **79**, 147–153 (2012).

[64] Catherine Winsper, Nicole K. Y. Tang, Steven Marwaha, Suzet Tanya Lereya, Melanie Gibbs, Andrew Thompson, and Swaran P. Singh, "The sleep phenotype of Borderline Personality Disorder: A systematic review and meta-analysis," Neuroscience & Biobehavioral Reviews **73**, 48–67 (2017).

[65] Silvia Miano, Ninfa Amato, Giuseppe Foderaro, Valdo Pezzoli, Gian Paolo Ramelli, Lorenzo Toffolet, and Mauro Manconi, "Sleep phenotypes in attention deficit hyperactivity disorder," Sleep Medicine **60**, 123–131 (2019).

[66] Justine M Kent, Sanjay J Mathew, and Jack M Gorman, "Molecular targets in the treatment of anxiety," Biological Psychiatry **52**, 1008–1030 (2002).

[67] Luc Staner, "Sleep and anxiety disorders," Dialogues in Clinical Neuroscience **5**, 249–258 (2003).

[68] Sarah L. Chellappa and Daniel Aeschbach, "Sleep and anxiety: From mechanisms to interventions," Sleep Medicine Reviews **61**, 101583 (2022).

[69] Amit Etkin and Tor D. Wager, "Functional Neuroimaging of Anxiety: A Meta-Analysis of Emotional Processing in PTSD, Social Anxiety Disorder, and Specific Phobia," American Journal of Psychiatry **164**, 1476–1488 (2007).

[70] Justin S. Feinstein, Dylan Gould, and Sahib S. Khalsa, "Amygdala-driven apnea and the chemoreceptive origin of anxiety," Biological Psychology **170**, 108305 (2022).

[71] Julian F. Thayer, Fredrik Åhs, Mats Fredrikson, John J. Sollers, and Tor D. Wager, "A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health," Neuroscience & Biobehavioral Reviews **36**, 747–756 (2012).

[72] Adrienne O'Neil, Shae E. Quirk, Siobhan Housden, Sharon L. Brennan, Lana J. Williams, Julie A. Pasco, Michael Berk, and Felice N. Jacka, "Relationship Between Diet and Mental Health in Children and Adolescents: A Systematic Review," American Journal of Public Health **104**, e31–e42 (2014).

[73] Piril Hepsomali and John A. Groeger, "Diet, Sleep, and Mental Health: Insights from the UK Biobank Study," Nutrients **13**, 2573 (2021).

[74] Irina Gaynanova, Naresh Punjabi, and Ciprian Crainiceanu, "Modeling continuous glucose monitoring (CGM) data during sleep," Biostatistics **23**, 223–239 (2022).

[75] María Arnoriaga-Rodríguez, Yenny Leal, Jordi Mayneris-Perxachs, Vicente Pérez-Brocal, Andrés Moya, Wifredo Ricart, Mercè Fernández-Balsells, and José Manuel Fernández-Real, "Gut Microbiota Composition and Functionality Are Associated With REM Sleep Duration and Continuous Glucose Levels," The Journal of Clinical Endocrinology & Metabolism , dgad258 (2023).

[76] Jessica A. Mong and Danielle M. Cusmano, "Sex differences in sleep: impact of biological sex and sex steroids," Philosophical Transactions of the Royal Society B: Biological Sciences **371**, 20150110 (2016).

[77] Eva Lindberg, Christer Janson, Thorarinn Gislason, Eythor Björnsson, Jerker Hetta, and Gunnar Boman, "Sleep Disturbances in a Young Adult Population: Can Gender Differences Be Explained by Differences in Psychological Status?" Sleep **20**, 381–387 (1997).

[78] Namni Goel, Hyungsoo Kim, and Raymund P. Lao, "Gender Differences in Polysomnographic Sleep in Young Healthy Sleepers," Chronobiology International **22**, 905–915 (2005).

[79] Ling-Ling Tsai and Sheng-Ping Li, "Sleep patterns in college students: Gender and grade differences," Journal of Psychosomatic Research **56**, 231–237 (2004).

[80] Ulrich Voderholzer, Anam Al-Shajlawi, Gesa Weske, Bernd Feige, and Dieter Riemann, "Are there gender differences in objective and subjective sleep measures? A study of insomniacs and healthy controls," Depression and Anxiety **17**, 162–172 (2003).

[81] Martica Hall, Raymond Vasko, Daniel Buysse, Hernando Ombao, Qingxia Chen, J. David Cashmere, David

Kupfer, and Julian F. Thayer, "Acute Stress Affects Heart Rate Variability During Sleep," Psychosomatic Medicine **66**, 56 (2004).

[82] Torbjörn Åkerstedt, Göran Kecklund, and John Axelsson, "Impaired sleep after bedtime stress and worries," Biological Psychology **76**, 170–173 (2007).

[83] Christopher L. Drake, Vivek Pillai, and Thomas Roth, "Stress and Sleep Reactivity: A Prospective Investigation of the Stress-Diathesis Model of Insomnia," Sleep **37**, 1295–1304 (2014).

[84] Shefali Liyanage, Kiran Saqib, Amber Fozia Khan, Tijhiana Rose Thobani, Wang-Choi Tang, Cameron B. Chiarot, Bara' Abdallah AlShurman, and Zahid Ahmad Butt, "Prevalence of Anxiety in University Students during the COVID-19 Pandemic: A Systematic Review," International Journal of Environmental Research and Public Health **19**, 62 (2022).

[85] Jane F. Gaultney, "The Prevalence of Sleep Disorders in College Students: Impact on Academic Performance," Journal of American College Health **59**, 91–97 (2010).

[86] Asma Ali Al Salmani, Asma Al Shidhani, Shatha Saud Al Qassabi, Shahad Ahmed Al Yaaribi, and Aysha Muslem Al Musharfi, "Prevalence of sleep disorders among university students and its impact on academic performance," International Journal of Adolescence and Youth **25**, 974–981 (2020).

[87] Chamara V. Senaratna, Jennifer L. Perret, Caroline J. Lodge, Adrian J. Lowe, Brittany E. Campbell, Melanie C. Matheson, Garun S. Hamilton, and Shyamali C. Dharmage, "Prevalence of obstructive sleep apnea in the general population: A systematic review," Sleep Medicine Reviews **34**, 70–81 (2017).

[88] Sirak Zenebe Gebreab, Caroline L. Vandeleur, Dominique Rudaz, Marie-Pierre F. Strippoli, Mehdi Gholam-Rezaee, Enrique Castelao, Aurélie M. Lasserre, Jennifer Glaus, Giorgio Pistis, Christine Kuehner, Roland von Känel, Pedro Marques-Vidal, Peter Vollenweider, and Martin Preisig, "Psychosocial Stress Over the Lifespan, Psychological Factors, and Cardiometabolic Risk in the Community," Psychosomatic Medicine **80**, 628 (2018).

[89] Johann Faouzi and Hicham Janati, "pyts: A Python Package for Time Series Classification," Journal of Machine Learning Research **21**, 1–6 (2020).

[90] Marc Maier, Hayley Carlotto, Sara Saperstein, Freddie Sanchez, Sherriff Balogun, and Sears Merritt, "Improving the Accuracy and Transparency of Underwriting with Artificial Intelligence to Transform the Life-Insurance Industry," AI Magazine **41**, 78–93 (2020).

[91] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, Vol. 5.1 (University of California Press, 1967) pp. 281–298.

[92] Saeed Aghabozorgi, Ali Seyed Shirkhorshidi, and Teh Ying Wah, "Time-series clustering – A decade review," Information Systems **53**, 16–38 (2015).

[93] Ali Javed, Byung Suk Lee, and Donna M. Rizzo, "A benchmark study on time series clustering," Machine Learning with Applications **1**, 100001 (2020).

[94] Romain Tavenard, Johann Faouzi, Gilles Vandewiele, Felix Divo, Guillaume Androz, Chester Holtz, Marie Payne, Roman Yurchak, Marc Rußwurm, Kushal Kolar, and Eli Woods, "Tslearn, A Machine Learning Toolkit for Time Series Data," Journal of Machine Learning Research **21**, 1–6 (2020).

[95] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 26 (1978) pp. 43–49.

[96] Zoltan Geler, Vladimir Kurbalija, Mirjana Ivanović, Miloš Radovanović, and Weihui Dai, "Dynamic Time Warping: Itakura vs Sakoe-Chiba," in *2019 IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)* (2019) pp. 1–6.

[97] Eshin Jolly, "Pymer4: Connecting R and Python for Linear Mixed Modeling," Journal of Open Source Software **3**, 862 (2018).

[98] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker, "Fitting Linear Mixed-Effects Models Using lme4," Journal of Statistical Software **67**, 1–48 (2015).

[99] Alexandra Kuznetsova, Per B. Brockhoff, and Rune H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," Journal of Statistical Software **82**, 1–26 (2017).