Objective assessment of sleep regularity in 60,000 UK Biobank participants using an open-source package

Daniel P. Windred[1], Samuel E. Jones[2,3], Alex Russell[1], Angus C. Burns[1], Philip Chan[1], Michael N. Weedon[2], Martin K. Rutter[4,5], Patrick Olivier[6], Céline Vetter[7], Richa Saxena[8,9], Jacqueline M. Lane[8,9], Sean W. Cain[1], and Andrew J. K. Phillips[1]

[1]Turner Institute for Brain and Mental Health, School of Psychological Sciences, Monash University, Melbourne, VIC, Australia

[2]Genetics of Complex Traits, University of Exeter Medical School, Exeter, UK

[3]Institute for Molecular Medicine Finland, Helsinki Institute of Life Science, University of Helsinki

[4]Centre for Biological Timing, Division of Endocrinology, Diabetes & Gastroenterology, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[5]Diabetes, Endocrinology and Metabolism Centre, Manchester University NHS Foundation Trust, Manchester, UK

[6]Action Lab, Department of Human-Centred Computing, Faculty of Information Technology, Monash University, Melbourne, VIC, Australia

[7]Circadian and Sleep Epidemiology Laboratory, Department of Integrative Physiology, University of Colorado Boulder, Boulder, CO, USA

[8]Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

[9]Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Corresponding author: Dr. Andrew Phillips

andrew.phillips@monash.edu

Monash University School of Psychological Sciences, Room 512, 18 Innovation Walk, Clayton Campus, VIC, 3800, Australia

Keywords

sleep habits; Sleep Regularity Index; intraindividual variability; circadian rhythm; circadian dysregulation; actigraphy; light; shift work; socioeconomic factors

Human health and behavior are regulated by a complex and extensive network of circadian clocks. These clocks are entrained by rhythmic signals in the environment, such as daily light exposure. In individuals who have irregular sleep schedules, these signals that furnish time-of-day information to the circadian system are less robust, which may cause circadian disruption and poor health outcomes.[1] Sleep regularity can be quantified using the Sleep Regularity Index (SRI),[2] a metric that compares sleep patterns between consecutive days (sleeping at similar times each day results in a high SRI). The SRI captures day-to-day variability in bedtime, waketime, sleep duration, naps, and awakenings during sleep.[3] Lower SRI has been associated with substantially increased risk for obesity, diabetes, cardiovascular disease, hypertension, and depressed mood.[4-6] Although sleep regularity is now recognized as a critical dimension of sleep health, there are barriers to measuring and reporting sleep regularity consistently. First, there are no open-source options for calculating sleep regularity, meaning it cannot be computed with the same ease as other common sleep metrics. Second, there is a lack of clear benchmarks for what represents a high or low level of sleep regularity at a population level, both for the SRI and other sleep regularity metrics.[3,4,6] We developed an open-source package for computing SRI from accelerometer data, and we applied it to the single largest accelerometer sample available to researchers, within the UK Biobank.

The UK Biobank is one of the largest and most comprehensive human health datasets.[7] In total, 103,104 participants (age (*M±SD*) = 62.3±7.9 years; 56.2% female) wore a wrist accelerometer (Axivity AX3[8]) to record daily rest-activity patterns for one week, between June 2013 and January 2016. We used the accelerometer data to estimate daily sleep onset and offset times with a validated, widely used R package ('GGIR'[9]). Since it is not possible to correctly calculate SRI from GGIR output alone, we developed our own accompanying open-source R package. Our package calculated sleep-wake state at the epoch-level and, unlike GGIR, it allowed multiple sleep bouts per day (allowing naps, fragmented sleep, or awakenings to be correctly factored into SRI calculation). The package also identified miscalculated sleep onset/offset times (5.7% of all nights in this dataset), which can lead to incorrect SRI scores (see Supplementary Material). After excluding participants with fewer than five days (120 hours) of 24h-separated epoch pairs containing valid sleep-wake data, the SRI was calculated in 60,997 participants (age (*M±SD*) = 62.7±7.8 years; 55.1% female).

The SRI scores had a median of 81.0 and interquartile range of 73.8-86.3 (*M±SD* = 78.8±10.7), shown in Figure 1A with a higher score indicating more regular sleep patterns. The distribution of SRI scores was non-normal with negative skew (KS test, *D*(60,997) = 0.99, *p*<.0001), consistent with distributions reported in smaller samples,[4] and 99% of individuals had SRI scores between 36.0-95.0. We observed a monotonic relationship between SRI and other common measures of sleep variability, including standard deviations in sleep onset, sleep offset, and sleep duration, shown in Figure 1B. These findings enable SRI scores to be related to equivalent values for other sleep variability measures, facilitating comparisons between studies that have reported different measures (see Supplementary Material). Across the sample, a cut-off of SRI<70 was comparable to *SD*>1.9 h for variability in sleep onset and offset, and *SD*>1.6 h for sleep duration. One in five individuals had an SRI below 71.6 (irregular), and one in five had an SRI above 87.3 (regular; Figure 1C).

Sleep regularity was related to several self-reported demographic variables, collected between 2006 and 2010. Lower SRI scores were found in those who were male (difference: -1.2, t-test, *p*<.0001), had higher material deprivation (Townsend Deprivation Index, bivariate correlation, *r*(60,922) = -0.11, *p*<.0001), had lower yearly household income (ANOVA, *p*<.0001), or had lower-level educational qualifications (ANOVA, *p*<.0001) (see Supplementary Material). Those of white ethnicity had significantly higher SRI than all other ethnic groups (2.6-6.8 points higher), and black ethnicity significantly lower (2.2-6.8 points lower; Kruskal-Wallis, *p*<.0001). Above the age of 65, sleep regularity decreased with age (bivariate correlation, *r*(27,918) = -0.02, *p* = 0.002). There was no relationship between age and SRI in people younger than 65 (*p* = 0.47), or across the whole sample age range (*p* = 0.13). Shift workers exhibited significantly lower SRI scores than non-shift workers (*M±SD* = 75.9±12.0 vs. 79.3±10.1, t-test, *p*<.0001). Employed people had significantly higher SRI than those who were Sick/Disabled, Unemployed, or Students (1.0-4.9 points higher), and lower SRI than Retired, Home/Family Caretakers, or Volunteers (0.2-1.6 points lower; Kruskal-Wallis, $p<.0001$). Together, these findings indicate that irregular sleep-wake patterns are associated with a complex set of individual and environmental factors, particularly socioeconomic disadvantage.

Sleep regularity is increasingly recognized as a fundamental determinant of health, and is a stronger predictor of cardiometabolic outcomes and quality of life than sleep duration.[4,10] The norms we begin to establish here provide a reference for researchers and clinicians intending to quantify sleep regularity with the SRI. Relative to other measures of variability in sleep timing, the SRI offers two key advantages: (i) it compares sleep-wake patterns between consecutive days (i.e., on a circadian timescale), meaning it may assess circadian disruption[3]; (ii) the SRI makes no assumptions about the structure of sleep (e.g., no assumption of one main sleep bout), making it applicable to populations with unusual sleep structure, such as shift workers or individuals with highly fragmented sleep.[1,11] We have developed a package to calculate SRI scores from accelerometer or binary sleep-wake data in R, available at https://github.com/dpwindred/sleepreg. The package can be used in combination with GGIR or as a standalone method for calculating SRI from pre-processed sleep-wake data. Processing of accelerometer data is based on the methods used here – estimating sleep-wake timing using GGIR, identifying fragmented sleep patterns, and identifying and excluding nights containing probable estimation errors. Our package also extracts percentile of calculated SRI scores, allowing other investigators' datasets to be compared to the UK Biobank SRI distribution, and generates raster plots of sleep-wake patterns. Given the widespread availability and use of consumer and research accelerometer data, our package will democratize the use of sleep regularity as an indicator of circadian health.

Disclosure Statement

Financial Disclosure: none.

Non-financial Disclosure: none.

References

1. Pan A, Schernhammer ES, Sun Q, Hu FB. Rotating night shift work and risk of type 2 diabetes: two prospective cohort studies in women. *PLoS Med*. 2011; 8 (12): e1001141.

2. Phillips AJK, Clerx WM, O'Brien CS, et al. Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing. *Sci Rep*. 2017; 7 (1): 3216.

3. Fischer D, Klerman EB, Phillips AJ. Measuring sleep regularity: Theoretical properties and practical usage of existing metrics. *Sleep*. 2021.

4. Lunsford-Avery JR, Engelhard MM, Navar AM, Kollins SH. Validation of the Sleep Regularity Index in Older Adults and Associations with Cardiometabolic Risk. *Sci Rep*. 2018; 8 (1): 14158.

5. Pye J, Phillips AJ, Cain SW, et al. Irregular sleep-wake patterns in older adults with current or remitted depression. *J Affect Disord*. 2021; 281: 431-437.

6. Fritz J, Phillips AJ, Hunt LC, et al. Cross-sectional and prospective associations between sleep regularity and metabolic health in the Hispanic Community Health Study/Study of Latinos. *Sleep*. 2021; 44 (4): zsaa218.

7. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015; 12 (3): e1001779.

8. Doherty A, Jackson D, Hammerla N, et al. Large scale population assessment of physical activity using wrist worn accelerometers: The UK Biobank Study. *PLoS One*. 2017; 12 (2): e0169649.

9. Migueles JH, Rowlands AV, Huber F, Sabia S, van Hees VT. GGIR: a research community–driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. *JMPB*. 2019; 2 (3): 188-196.

10. Trivedi R, Man H, Madut A, et al. Irregular sleep/wake patterns are associated with reduced quality of life in post-treatment cancer patients: a study across three cancer cohorts. *Front Neurosci*. 2021.

11. Senaratna CV, Perret JL, Lodge CJ, et al. Prevalence of obstructive sleep apnea in the general population: a systematic review. *Sleep Med Rev*. 2017; 34: 70-81.

Figure 1. (A) Distribution of Sleep Regularity Index scores in 60,997 UK Biobank participants, grouped by percentile ranges. Inset raster plots of sleep-wake patterns in three participants represent regular (SRI = 94; top), slightly irregular (SRI = 70; middle), and highly irregular (SRI = 34; bottom) sleep-wake patterns. (B) Mean intra-individual standard deviation in sleep timing and duration for each percentile range [colour-matched to (A)], calculated using each participant's one-week data collection period. (C) Comparative distributions of standard deviation in sleep timing and duration for upper (SRI>87.3) and lower (SRI<71.6) quintiles.

Supplementary Material

1. Sleep Regularity Index (SRI)

SRI scores were calculated using the following equation:

$$SRI = -100 + 200\left(1 - \frac{1}{N_v}\sum_{i=1}^{N}|s_i - s_{i+C}|\right)$$

Sleep-wake state is represented by $s_i = 1$ for wake, $s_i = 0$ for sleep, and $s_i = NA$ represents excluded epochs. Number of valid epoch-by-epoch comparisons is represented by N_v, which includes all comparisons where $s_i \neq NA$ and $s_{i+c} \neq NA$. Where $s_i = NA$ or $s_{i+c} = NA$, $|s_i - s_{i+c}| = 0$. Subscript i represents each epoch from recording start to 24h prior to recording end, such that at:

$i = 1, \quad t_1 = 0$

$i = 2, \quad t_2 = E$

$i = 3, \quad t_3 = 2E$

$\vdots$

$i = C, \quad t_C = E(C-1) = 24$

$\vdots$

$i = N, \quad t_N = E(N-1) = t_{max} - 24$

Time is represented by t_i, epoch length is represented by E, recording length is represented by t_{max}, and number of epochs within one 24-h interval is represented by C. All time values are in hours.

An SRI score of 100 represents perfectly regular day-to-day sleep-wake patterns, and an SRI score of 0 represents random patterns. While the possible range of SRI scores is -100 to 100, scores rarely fall below zero in practice. A minimum of five days (120 hours) of overlapping valid epochs are required for SRI calculation. This is recommended to reduce variability caused by single days of irregular sleep timing in short samples [1]. Missing epochs of data are excluded in the SRI calculation. The five-day minimum rule and handling of missing epochs are built into our package.

We recommend using at least 7 consecutive days of sleep-wake data to calculate SRI scores. This will minimize violations of the five day overlap rule, while also capturing weekday-weekend variability in sleep-wake patterns.

2. Identification of daily sleep onset and offset timing with GGIR

Daily sleep-wake timing predictions from accelerometer data were made using GGIR (version 2.0-0).[2,3] A full description of the GGIR package can be found at: https://cran.r-project.org/web/packages/GGIR/GGIR.pdf.

Raw data from Axivity AX3 devices were converted to .csv format and downsampled to 1 Hz to speed up computation for the very large dataset, before being input to the 'read.myacc.csv' function in GGIR, which extracts accelerometer data from .csv files. A full list of our specified GGIR parameters can be found in our package source code, and in GGIR's 'config' file output. One noteworthy non-default parameter was the window size for detecting device angle change, which we increased from 5 seconds to 15 seconds, allowing for more accurate detection of sustained inactivity and sleep-wake times in our downsampled (1 Hz) data ('windowsize' parameter[3]).

GGIR's 'sleeponset' and 'wakeup' variables were used as daily sleep onset and offset times, and 'SleepDurationInSpt' was used as sleep duration (representing sustained inactivity time within GGIR's daily 'sleep window'). Standard deviations were calculated at the intra-individual level across all valid days of actigraphy data. GGIR defined invalid days as any interval from noon-noon with >4-h of data classified as missing, clipped, or non-wear.

3. Identification of naps and fragmented sleep patterns

GGIR assumes only one sleep episode per day as one of its key rules. Since naps and fragmented sleep patterns contribute substantially to SRI scores, we accounted for these in our SRI calculation.

GGIR defines wake after sleep onset (WASO) as any epoch within each 'sleep window' that does not contain sustained accelerometer inactivity. We defined 'WASO' as any epoch of 30 minutes or longer without any sustained inactivity, inside GGIR-defined sleep windows. We defined 'naps' as any interval of 30 minutes or longer with at least 95% sustained inactivity (excluding intervals classified as non-wear time, described below), outside GGIR-defined sleep windows. Sustained inactivity is defined by GGIR as any continuous period of 5 minutes or longer where the angle of the device relative to the z-axis does not change by more than 5 degrees.

4. Identification of miscalculated sleep onset and offset times

We chose to identify and exclude days where sleep onset or offset times were likely miscalculated by GGIR, after identification of such cases by visual inspection across a subset of the data. Miscalculated times were identified based on sustained inactivity within the 1.5-h intervals both before and after onset and offset of GGIR's predicted daily sleep windows. For the two 1.5-h intervals inside each sleep window, onset/offset times were classified as miscalculated if <20% of the interval was sustained inactivity (representing probable wake where sleep should be present). Similarly, for intervals outside the sleep window, onset/offset times were classified as miscalculated if >85% of the interval was sustained inactivity (representing probable sleep where wake should be present). Percentages values were determined by visual inspection of accelerometer data and predicted onset/offset times.

Classification of miscalculated days is optional in our package, since sleep-wake timing predictions may improve in subsequent versions of GGIR.

5. Non-wear detection

Non-wear detection was implemented according to methods described in van Hees et. al.[4] In addition to GGIR's built-in non-wear detection algorithm, we also wrote our own version of this method, allowing for customizability of parameters, output of non-wear data as .csv files, and to catch cases where GGIR's non-wear detection failed.

All identified non-wear epochs outside GGIR's sleep windows were excluded prior to SRI calculation. Identified non-wear epochs within GGIR's sleep windows were not excluded, due to the tendency of the algorithm to identify short intermittent periods of non-wear during sleep (i.e., when a participant is very still, but likely wearing the device).

All days with more than 6 h of non-wear were excluded, to prevent cases of systematic non-wear introducing bias in SRI scores (e.g., removing the device for the duration of the sleep period every night).

6. 'sleepreg' package

Our package is available for download via Github: https://github.com/dpwindred/sleepreg

Calculate SRI from accelerometer data

Downsampling ('ds_accel_csv'). Accelerometer files are downsampled to 1 Hz creating a consistent input format for GGIR, and increasing speed. Recording frequency of greater than or equal to 1 Hz is required for input files.

Non-wear detection ('nonwear_detect'). Non-wear is evaluated in 15-minute epochs, based on surrounding 60-minute windows. Standard deviation <13mg and range <50mg in at least two accelerometer axes is required for non-wear classification.

GGIR ('GGIR_from_csv'). Specifies parameters and implements GGIR across .csv accelerometer files, extracting sleep-wake predictions and sustained inactivity bouts.

Calculate SRI from GGIR output ('SRI_from_GGIR'). Uses sleep windows and sustained inactivity bouts from GGIR output to calculate SRI scores. Accounts for multiphasic and fragmented sleep and excludes days where sleep onset and offset times are miscalculated, using methods described here. Runs across GGIR output directories, accounting for both multi-file and single-file GGIR output structures. Additional outputs include sleep-wake raster plots, summary of miscalculated nights, and binary sleep-wake vectors.

Wrapper ('SRI_from_accel_csv'). Sequentially runs downsampling, non-wear detection, GGIR, and SRI calculation.

Calculate SRI from sleep diary or other binary sleep-wake data

'SRI_from_binary'. SRI scores are calculated from binary sleep-wake data in required input format, alongside sleep-wake raster plots.

7. SRI and intra-individual variability in sleep onset, offset, and duration

Table S1. SRI scores and associated intra-individual variability in sleep onset, offset, and duration across one-week actigraphy recordings. Onset and offset are defined by GGIR's 'sleeponset' and 'wakeup' variables. Duration is defined by GGIR's 'SleepDurationInSpt' variable, which represents total sustained inactivity time between daily 'sleeponset' and 'wakeup' times (i.e., duration = offset – onset – wake after sleep onset).

		Average Standard Deviation (h)		
SRI Range	Percentile Range	Sleep Onset	Sleep Offset	Sleep Duration
<46	0.0%-1.6%	2.64	2.61	1.73
46-48	1.6%-1.9%	2.23	2.25	1.77
48-50	1.9%-2.3%	2.50	2.32	1.74
50-52	2.3%-2.7%	2.44	2.33	1.80
52-54	2.7%-3.2%	2.42	2.27	1.85
54-56	3.2%-3.9%	2.28	2.26	1.75
56-58	3.9%-4.8%	2.18	1.99	1.73
58-60	4.8%-5.9%	2.06	1.92	1.66
60-62	5.9%-7.2%	1.97	1.85	1.67
62-64	7.2%-8.9%	1.86	1.69	1.62
64-66	8.9%-11.0%	1.74	1.68	1.58
66-68	11.0%-13.6%	1.70	1.59	1.56
68-70	13.6%-16.8%	1.58	1.53	1.52
70-72	16.8%-20.8%	1.47	1.40	1.43
72-74	20.8%-25.5%	1.39	1.32	1.38
74-76	25.5%-31.2%	1.27	1.24	1.31
76-78	31.2%-37.9%	1.20	1.18	1.27
78-80	37.9%-45.7%	1.09	1.11	1.21
80-82	45.7%-54.5%	1.00	1.02	1.13
82-84	54.5%-64.2%	0.91	0.98	1.06

84-86	64.2%-73.7%	0.84	0.91	1.00
86-88	73.7%-82.9%	0.73	0.84	0.90
88-90	82.9%-90.7%	0.66	0.78	0.85
90-92	90.7%-95.9%	0.58	0.71	0.77
92-94	95.9%-98.8%	0.51	0.62	0.72
94-96	98.8%-99.8%	0.41	0.49	0.62
>96	99.8%-100.0%	0.41	0.42	0.56

8. Relationships of SRI with demographic variables

All demographic variables were collected during participants' initial assessment visit (2006-2010), with the exception of sex, date of birth, and Townsend Deprivation Index, which were based on NHS Registry data and updated during initial assessment if required. Accelerometer data used to calculate SRI scores was collected between June 2013 and January 2016. Participants were instructed to wear devices continuously for seven days under free-living conditions. Complete participant instructions are available at: https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=141141

Ethical approval was granted by the North West Multi-centre Research Ethics Committee (MREC), covering the UK. Approvals have also been granted by the National Information Governance Board for Health & Social Care (NIGB) in England and Wales, and the Community Health Index Advisory Group (CHIAG) in Scotland. Informed consent was obtained from all participants.

Sex (Field ID: 31)

There was a significant difference in SRI scores between males ($M \pm SD$ = 78.1±11.1) and females ($M \pm SD$ = 79.3±10.3) (two-sample t-test, $t(56,788) = 13.6$, $p<.0001$, Cohen's $d = 0.11$).

Age

Age at actigraphy commencement was calculated based on year and month of birth in 60,993 participants.

Qualifications (Field ID: 6138)

Six educational qualifications were compared: Certificate of Secondary Education (CSE), National Vocational Qualification (NVQ) / Higher National Diploma (HND) / Higher National Certificate (HNC), Other qualifications (e.g., nursing, teaching), O Levels (completion of school years 10-11), A Levels (completion of secondary education, school years 12-13;), and University. A total of 55,153 had provided their qualifications and had valid SRI scores.

There was a significant difference in SRI scores between yearly household income groups (ANOVA, $F(5, 55{,}147) = 39.3$, $p<.0001$). Multiple comparisons (Tukey's HSD) revealed those with higher levels of educational qualifications also exhibited higher SRI scores (all with $p<.01$), with the exception of the three relationships between O Levels, A Levels, and Other qualifications (e.g., nursing, teaching), and the relationship between CSE (pre-O Level) and NVQ/HND/HNC (vocational) qualifications, which were non-significant (Figure S1). For example, those with University-level qualifications had higher SRI scores than all other groups, and those with CSE qualifications had lower SRI than all other groups except for those with NVQ/HND/HNC qualifications.

Yearly household income (Field ID: 738)

Yearly household income was compared across five income brackets: <£18k, £18k-£29.9k, £30k-£51.9k, £52k-£100k, >£100k. A total of 54,621 provided their yearly household income and had valid SRI scores.

There was a significant difference in SRI scores between yearly household income groups (ANOVA, $F(4, 54{,}616) = 139.4$, $p<.0001$). Multiple comparisons (Tukey's HSD) revealed significantly lower SRI scores in groups with lower yearly income (all with $p<.0001$), with the exception of the relationship between the two highest income levels, which was non-significant (Figure S1).

Ethnic background (Field ID: 21000)

Ethnic background was collected under six broad categories: White (97.2%), Asian (0.9%), Black (0.7%), Chinese (0.2%), Mixed (0.5%), and Other (0.5%). A total of 60,780 provided their ethnic background and had valid SRI scores.

There was a significant difference in SRI scores between ethnic groups (Kruskal-Wallis test, $H(5) = 331.68$, $p<.0001$). Multiple comparisons (Dunn's test, Benjamini-Hochberg) revealed significantly higher SRI scores in those of white ethnicity (2.6-6.8 points higher), and significantly lower SRI scores in those of black ethnicity (2.2-6.8 points lower). There were no significant differences between any other ethnic groups (Figure S1).

Townsend Deprivation Index (Field ID: 189)

The Townsend Deprivation Index, a measure of material deprivation encompassing unemployment, home ownership, car ownership, and household overcrowding, was available in 60,924 participants who also had valid SRI scores.

Employment status (Field ID: 6142)

Employment status categories were Employed (60.8%), Retired (33.1%), Home/Family Caretaker (2.8%), Sick/Disabled (1.5%), Unemployed (1.0%), Volunteer (0.5%), and Student (0.2%). A total of 60,595 provided their employment status and had valid SRI scores.

There was a significant difference in SRI scores between employment status groups (Kruskal-Wallis test, $H(6) = 329.4$, $p<.0001$). Multiple comparisons (Dunn's test, Benjamini-Hochberg) revealed significant differences between all pairs except for Unemployed-Student, Retired-Volunteer, and Home/Family Caretaker-Volunteer (Figure S1).

Shift work status (Field ID: 826)

Participants were asked whether their job involved shift work, and answered based on four categories: Never / rarely, Sometimes, Usually, and Always. We classified people responding Never / rarely as non-shift workers (*N* = 32,353), and people responding with Sometimes, Usually, or Always as shift workers (*N* = 4,435).

There was a significant difference in SRI scores between shift workers ($M \pm SD$ = 75.9±12.0) and non-shift workers ($M \pm SD$ = 79.3±10.1) (two-sample t-test, $t(5332.8) = 18.1$, $p<.0001$, Cohen's $d = 0.31$).

References

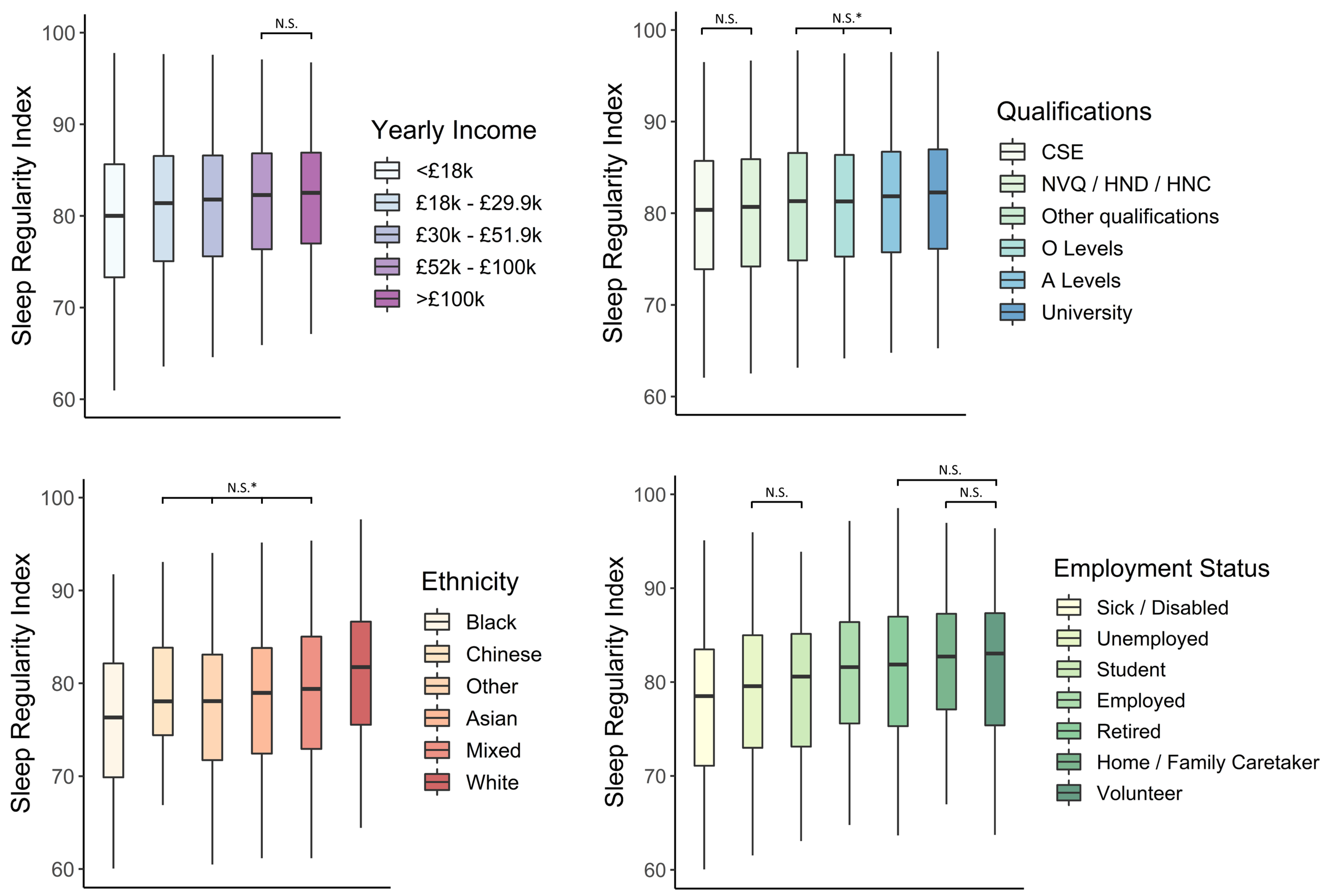
1. Fischer D, Klerman EB, Phillips AJ. Measuring sleep regularity: Theoretical properties and practical usage of existing metrics. *Sleep*. 2021.

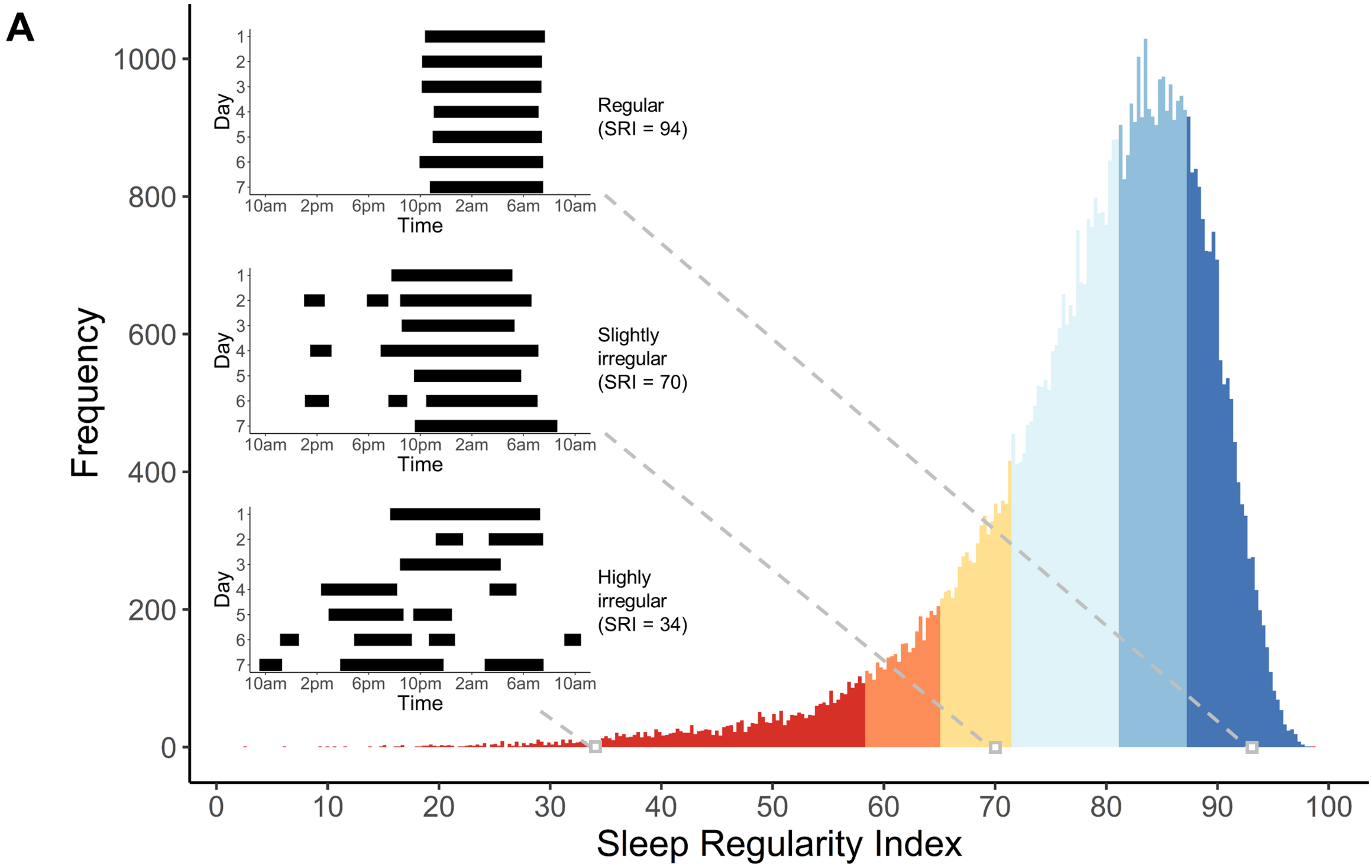
2. Migueles JH, Rowlands AV, Huber F, Sabia S, van Hees VT. GGIR: a research community–driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. *JMPB*. 2019; 2 (3): 188-196.

3. van Hees VT, Sabia S, Jones SE, et al. Estimating sleep parameters using an accelerometer without sleep diary. *Sci Rep*. 2018; 8 (1): 1-11.

4. van Hees VT, Renström F, Wright A, et al. Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. *PLoS One*. 2011; 6 (7): e22922.

Figure S1. SRI scores across yearly income, educational qualification, ethnic background, and employment status categories (median, IQR). In each figure, post-hoc comparisons revealed significant pairwise differences between all group except those labelled as non-significant (N.S.). Categories are ordered by median SRI score, from lowest to highest. *There were no significant pairwise differences between any of these groups.


Sleep Regularity Index
100
90
80
70
60
N.S.
Yearly Income
<£18k
£18k - £29.9k
£30k - £51.9k
£52k - £100k
>£100k
Sleep Regularity Index
N.S.
N.S.*
Qualifications
CSE
NVQ / HND / HNC
Other qualifications
O Levels
A Levels
University
Sleep Regularity Index
N.S.*
Ethnicity
Black
Chinese
Other
Asian
Mixed
White
Sleep Regularity Index
N.S.
N.S.
N.S.
Employment Status
Sick / Disabled
Unemployed
Student
Employed
Retired
Home / Family Caretaker
Volunteer



B

Percentile	SRI	Standard Deviation (h) Sleep Onset	Sleep Offset	Sleep Duration
0-5%	< 58.4	2.41	2.32	1.75
5-10%	58.4 - 65.1	1.90	1.78	1.63
10-20%	65.1 - 71.6	1.59	1.51	1.51
20-50%	71.6 - 81.0	1.19	1.18	1.27
50-80%	81.0 - 87.3	0.86	0.93	1.02
80-100%	87.3 - 100	0.61	0.73	0.80