Objectively measured sleep duration, sleep quality, and symptoms of psychopathology in the SIGMA cohort

Noëmi Hagemann^{1, 2*}, Olivia J. Kirtley¹, Ginette Lafit^{1,3}, Martien Wampers¹, Robin Achterhof¹, Karlijn F.S.M Hermans¹, Anu P. Hiekkaranta¹, Aleksandra Lecei⁴, Davy Vancampfort² & Inez Myin-Germeys¹

¹ *KU Leuven, Department of Neurosciences, Research Group Psychiatry, Center for Contextual Psychiatry, Leuven, Belgium*

² *KU Leuven, Department of Rehabilitation Sciences, Research Group Adapted Physical Activity and Psychomotor Rehabilitation, Leuven, Belgium*

³ *KU Leuven, Department of Psychology, Research Group on Quantitative Psychology and Individual Differences, Leuven, Belgium*

⁴ *KU Leuven, Department of Neurosciences, Research Group Psychiatry, Center for Clinical Psychiatry, Leuven, Belgium*

*Corresponding author: Noëmi Hagemann, Department of Neurosciences, Research Group Psychiatry, Center for Contextual Psychiatry, Kapucijnenvoer 33 bus 7001, 3000 Leuven, Belgium

E-mail: noemi.hagemann@kuleuven.be

Abstract

While subjective measures have demonstrated an association between sleep duration, sleep quality, and symptoms of psychopathology in adolescents, findings from more reliable, objective measures remain limited and inconsistent. In this study, we investigate if objectively measured sleep duration and sleep quality are associated with symptoms of psychopathology in adolescents.

Five hundred-fifty-eight adolescents aged 11-17 years from the SIGMA cohort wore the Fitbit Charge 2 measuring sleep duration and sleep quality during 4-6 days. Participants completed the Brief Symptom Inventory-53, assessing symptoms of general psychopathology, depression, anxiety, and psychoticism. Power analysis and non-linear model selection were conducted within a training subset (n=56). The optimal non-linear and linear models were fit to the remainder of the dataset (n=502).

Neither short nor long sleep duration were associated with symptoms of general psychopathology, depression, anxiety, and psychoticism in the overall sample and in younger (11-14-year-old) and older (15-17-year-old) adolescents. Further, better sleep quality was not associated with symptom levels of anxiety, and psychoticism, but small to negligible positive associations were found between sleep quality and general psychopathology symptoms and symptoms of depression.

Our non-significant findings converge with those of an increasing body of objectively measured sleep literature that does not find significant associations between sleep duration, sleep quality and psychopathology symptoms. Overall, our results suggest that associations between sleep duration and psychopathology symptoms in previous studies may be a function of the subjective, self-report nature of the employed measures, and do not generalize to objectively collected sleep data.

Keywords: sleep duration; sleep quality; psychopathology; adolescents; accelerometry

Data availability statement

Data are available upon request from Olivia Kirtley [olivia.kirtley@kuleuven.be] and Inez Myin-Germeys [inez.germeys@kuleuven.be]

Acknowledgements

We would like to thank the research assistants Tessa Biesemans, Lore Depraetere and Jolien Bynens, data manager Martien Wampers, research coordinator Silke Apers, and administrative and financial coordinator Martine van Nierop for their invaluable work. We are extremely grateful for the assistance from master students Louise Dendoncker, Lore Dumolein, Romanie Kint, and Silke De Neve in the data extraction process. We would also like to thank the collaborating schools, teachers, and participants from the SIGMA study, who were incredibly supportive in the SIGMA study collaboration.

Funding

Noëmi Hagemann, Ginette Lafit, Robin Achterhof, Karlijn Hermans, Anu Hiekkaranta, Martien Wampers, and Inez Myin-Germeys are supported by a Research Foundation Flanders Odysseus grant to Inez Myin-Germeys (FWO G0F8416N). Noëmi Hagemann is also supported by a doctoral grant by the German scholarship foundation Cusanuswerk e.V. Olivia Kirtley is supported by a Senior Postdoctoral Research Fellowship (FWO 1257821N) from Research Foundation Flanders. Aleksandra Lecei is supported by a Doctoral Grant (FWO 1104219N) from Research Foundation Flanders.

Conflicts of interest/Competing interests

The authors have no financial or non-financial interests to disclose.

Introduction

Adolescence is a vulnerable developmental period for the onset of psychopathology symptoms (Patton et al., 2014), and 50% of these mental health symptoms persist into early adulthood (Kessler et al., n.d.). Two risk factors that may increase symptoms of psychopathology, are inadequate sleep duration and quality. Adequate sleep refers to age-specific sleep duration needs, while good sleep quality is characterized by continuous, uninterrupted sleep, and feelings of refreshment upon awakening (Kline, 2013). Current literature suggests that adequate sleep duration and good quality sleep are essential for adolescents' mental and physical health (Cairns et al., 2014; Matricciani et al., 2019). Most guidelines, such as from the American Society of Sleep Medicine and the National Sleep Foundation, contend that younger adolescents between 11-13 years old should sleep 9-11 hours, and older adolescents between 14-17 years old should sleep between 8-10 hours per night, to maintain general health (Paruthi et al., 2016). Yet, the percentage of adolescents reaching the recommended sleep duration is low. Only about two thirds of the younger and older adolescents sleep the recommended duration of at least 8-9 hours per night on a school night, according to international studies (Delaruelle et al., 2019; Michaud & Chaput, 2016).

Previous longitudinal studies demonstrate that both very short and very long sleep durations are associated with subsequent psychopathology symptoms (Fitzgerald et al., 2011; Kaneita et al., 2007; Liu et al., 2020; Morishima et al., 2020; Ojio et al., 2016). Sleeping shorter or longer than recommended by guidelines (i.e. 8-9 hrs per night) on a school night is associated with a higher risk of developing symptoms of depression and anxiety (Kaneita et al., 2007; Liu et al., 2020; Ojio et al., 2016). Overall, as findings show associations of both short and long sleep duration with psychopathology symptoms, a non-linear (U-shaped) relationship between sleep duration and psychopathology has been suggested (Liu et al., 2020). Furthermore, prospective studies provide stronger evidence for sleep duration predicting the development of psychopathology rather than vice versa, however few studies have investigated the potential longitudinal effects of depressive symptoms on sleep duration (Lovato & Gradisar, 2014). In contrast, the prospective relationships between lower sleep quality and more symptoms of depression and anxiety is considered bi-directional (Kaneita et al., 2009).

Despite this seemingly convincing literature, no consensus exists about whether sleep relates to psychopathology. There is a great discrepancy between the literature mentioned above, which relies on self-report data, and recent studies applying objective i.e. non-self-report, measures to investigate the sleep – psychopathology relationship. In fact, the small number of objective data studies investigating the longitudinal association between sleep duration, symptoms of general psychopathology (Ranum et al., 2019) and depression (Slykerman et al., 2019), do not find evidence for a sleep duration – psychopathology relationship. Similarly, recent research found no association between objectively measured sleep quality and symptoms of psychoticism (Koopman-Verhoeff et al., 2019). Overall, these objective findings are meaningful, since they challenge the conclusions from previous literature, which is largely based on retrospective self-report.

Self-report data are considered less reliable, as they overestimate sleep duration by 30-60 minutes compared to actigraphy and polysomnography (Guedes et al., 2016; Matthews et al., 2018). Research devices (e.g. actigraphy and polysomnography) are the best tools for accurately measuring sleep, but one downside is that they are relatively expensive – thereby limiting the sample sizes for studies that use these methods. However, more cost-effective alternatives such as multi-sensor consumer devices measure sleep duration in adolescents with comparable accuracy to actigraphy (Lee et al., 2019; Menghini et al., 2021) and provide almost equally accurate sleep epochs, such as total sleep time (i.e. sensitivity (0.95-0.96) and specificity (0.58-0.69)) when compared to polysomnography in adults (Haghighat et al., 2019). Whilst multi-sensor consumer devices cannot replace polysomnography, their cost-effectiveness facilitates collection of objective sleep data from larger samples.

Previous literature has not only focused mainly on self-report measures, but was also limited to specific psychopathologies such as depression, anxiety, and psychoticism. In contrast, broader associations with general psychopathology symptoms remain unknown. This is particularly relevant in adolescence, when psychopathology is still developing and less differentiated than in adults (Hartmann et al., 2021). More research is necessary in order to better understand how sleep duration and quality relate to general psychopathology. Yet, since studies using objective measures are still limited, exploring the relationship of sleep duration and quality with symptoms of depression, anxiety, and psychoticism remains relevant as well. Moreover, no previous study has investigated a non-linear

relationship between sleep duration and psychopathology in objectively measured data, despite both short and long sleep duration being associated with psychopathology symptoms (Liu et al., 2020).

In order to disentangle the nature of the association between objectively measured sleep duration and symptoms of general psychopathology, and to understand whether complying with current guidelines is associated with less psychopathology in adolescents, we investigate if there is a non-linear association between multisensory-device derived sleep duration and symptoms of psychopathology. Further, we assess whether objective sleep quality is associated with symptoms of psychopathology. For these post-registered¹, confirmatory analyses, we hypothesize that there is a non-linear association between sleep duration and symptoms of psychopathology where low and high sleep duration are associated with more psychopathology, and a moderate sleep duration is associated with the least psychopathology (H1). Further, we hypothesise that non-compliance with the age-specific guidelines for sleep duration in younger (11-14-year-old) and older adolescents (15-17-year-old) will be associated with more symptoms of psychopathology by adding interaction effects between sleep duration and age group (i.e. younger [11-14-year-old] and older adolescents [15-17-year-old]) (H2). Lastly, we hypothesize that a linear association exists between sleep quality and psychopathology (H3).

Subsequently, we conducted exploratory (non-registered) analyses to examine depression, anxiety, and psychoticism symptoms and their relationships to sleep duration, quality, and adherence to sleep guidelines to compare our findings with the previous literature.

Methods

Sample and recruitment

Data from Wave 1 of the SIGMA study, a large-scale, ongoing, accelerated longitudinal study of adolescent mental health in Flanders, Belgium, are used in this study. In total, 1913 adolescents were included in Wave 1 (age range 11-19 at enrolment). Full details of the methods, measures and sample have been described elsewhere (Kirtley et al., 2021). The participants of the SIGMA study

¹ A post-registration is a form of pre-registration that occurs following data collection, but before data access and analysis (Benning et al., 2019)

were recruited as a general population sample via schools, with an opt-in consent procedure. Students at each school were invited to take part in the study, regardless of sociodemographic or psychological factors, e.g. known presence/absence of psychological disorder. The inclusion criteria were: (a) being a current student in the first, third or fifth year at a participating school; and (b) being able to read and write Dutch to an acceptable level. Ethical approval for this study was obtained from the UZ/KU Leuven Medical Ethics Committee (Ref: S6 1395).

Procedure

Self-reported symptoms of psychopathology were assessed via a tablet computer, in class groups at school. Following completion, students were instructed to wear a Fitbit Charge 2 device on their wrist, during 7 consecutive days and 24 hours per day. At night, the Fitbit Charge 2 automatically collects data on the participant's sleep duration and sleep quality. Participants were asked to remove the device whenever they would be in contact with water (e.g. when swimming, showering) since the device is not water resistant. In addition to wearing the Fitbit device, participants were also prompted to answer ten brief questionnaires per day for six days via smartphone during 7:30 am to 22:30 pm as part of the experience sampling method (ESM) component of the larger study (Myin-Germeys et al., 2018). These ESM data were not used within the current study. To avoid conflict between the ESM questionnaires and participants' sleep schedules, we instructed them to follow their normal individual daily life rhythm, such as sleeping in during the weekend or going to bed early, even if this meant missing ESM questionnaires. We also ensured that the measurement week was a usual week of participants' daily lives to ensure that the sleep variables would be representable. After the 7-day Fitbit wear period, they returned the material at their school. Participants received a 10-euro voucher for participation.

Symptoms of psychopathology

The Brief Symptom Inventory – 53 (BSI-53; Derogatis, 1993) was used to assess symptoms of psychopathology. It includes 10 different subscales on somatization, compulsiveness, social insecurity, depression, anxiety, aggression, phobia, paranoia, psychoticism, and additional items (e.g. suicidal thoughts, sleep quality). Example items are feeling unworthy, nervous, or distrustful to most people from the depression, anxiety, and psychosis subscales respectively. Participants were asked to report on all 53 items how much they had experienced them during the last week, including the day of assessment on a scale ranging from 0 ('not at all') to 4 ('very much'). They also had a response option 'I do not want to answer this question'.² We used the Global Severity Index (GSI), which is a mean score of all 53 items and also mean values of the subscales of depression, anxiety, and psychosis. One of the 53 items is on sleep quality. Due to the overlap with our predictor variable, this variable was excluded, and the BSI-GSI value was calculated based on 52 items. The Dutch version of the BSI-53 is sufficiently valid and reliable (Beurs & Zitman, 2006). In the current study, overall reliability was good: McDonald's Omega $\omega = .97$, as was the reliability of the subscales depression ($\omega = .90$), anxiety ($\omega = .87$), and psychoticism ($\omega = .75$). GSI scores were calculated in accordance with the BSI-53 guidelines. Therefore, participants with BSI- 53 data containing more than 3 missing values in total or more than 1 missing value per subscale were excluded.

Sleep

We measured sleep duration and sleep quality with the consumer tracker device Fitbit Charge 2. Sleep duration is considered as the time spent asleep, and therefore automatically excludes periods in bed which had been recorded as being awake or restless during the tracked sleep time (Fitbit, 2021a). Sleep quality is assessed by a sleep quality score which is based on heart rate, the time spent awake or restless, and sleep stages. This score is based on a proprietary algorithm, meaning that information about the exact method used to calculate the sleep quality score is not publicly available. The score ranges from 0-100, with a higher score indicating a higher level of sleep quality. A sleep score of 90-100 is defined by Fitbit as 'excellent', 80-89 as 'good', 60-79 as 'fair', and a score below 60 as 'poor' (Fitbit,

² This option was introduced to all questionnaires at the request of schools (see Kirtley et al., in preparation, for further information).

2021b). The screen of the device was disabled to prevent participants receiving feedback their activity (which may influence their behaviour), but the distance (km) they covered daily was still displayed as this could not be turned off. To create average minutes of sleep duration and sleep quality, the sum of sleep minutes and the sum of the sleep quality score over the whole valid wear time period were divided by the number of valid measurement days.

Inclusion and exclusion of data in the final dataset

We used a threshold of at least ≥ 4 valid measurement days as an inclusion criterion in line with previous research (Migueles et al., 2017). A valid day was defined as having a measurement day with a sleep duration of ≥ 200 minutes and where sleep started after 9 pm. For more details regarding the data pre-processing steps and in- and exclusion criteria, please see **Supplement 1**.

Covariates

Meta-analytic evidence suggests that adolescents' sleep duration decreases with older age (Galland et al., 2018), and that girls tend to sleep longer and have better sleep quality (Franco et al., 2020). Further, a cross-sectional study from Germany with comparable seasonal conditions shows that adolescents sleep less (~ 40 min) in summer as compared to winter, and that warm weather reduces sleep quality (Quante et al., 2019). Therefore, we included age (years), gender (male/female), and season (spring/summer/autumn/winter) as covariates into our analyses.

Open Science Practices

This study was post-registered - (a form of pre-registration that occurs following data collection, but prior to data access and analyses (Benning et al., 2019) - on the Open Science Framework (see: https://osf.io/8v3m5/?view_only=b023aa5d3af948a6919a73cf2902254e). Deviations from the post-registration are reported in **Supplement 2**.

Statistical analysis

The aims of the statistical analysis were threefold. Given that we wanted to assess the non-linear relationship between sleep duration and psychopathology symptoms, we used a linear regression model including cubic regression splines, which enable us to investigate the non-linear relationship. Cubic regression splines have a high degree of freedom, and are therefore able to catch the variability

of slope flexibly. Also, they allow smooth interpolation of fixed points, called knots (Harrell, 2020). In this study, the use of cubic splines allows us to accurately determine the potential ‘cut-off points’ where sleep duration becomes (mal)adaptive’.

To prepare for our analyses, we first assessed how many cubic spline knots provided the best fit. Second, after the best fitting model had been determined, we conducted a two-step simulation-based power analysis to understand if our sample was sufficiently powered. To address these two aims, we conducted a few preparatory steps; (a) a data cleaning procedure according to the pre-defined in- and exclusion criteria; and subsequently, (b) a sample selection, where we drew a random stratified 10% subsample considering age and gender from the total, cleaned sample. After we imputed the missing data within this 10% training subsample, we, first, compared various numbers of knots in cubic splines and determined which linear mixed including either, 3, 4, or 5 knots provided the best fit. Second, after the model was specified, we used the 10% training subsample to conduct two separate simulation-based power analyses. Third, and finally, we estimated the models within remaining dataset, the 90% testing dataset.

Preparatory steps

First, we applied the inclusion and exclusion criteria to the full dataset (N=1178). We performed this step before dividing our sample into training and testing samples, as we were aware of a high amount of missing sleep data (~58%) from a previous study (Hagemann et al., 2021). After having excluded participants (a) whose device did not record any data during the complete 8-day measurement period (N= 247), (b) were older than 17 years of age (N=26), and (c) had less than 4 days of valid data (N= 347), the total sample consisted of 558 participants. Within the cleaned sample, 32.9% of the observations within the variables sleep duration and sleep quality were missing. We used Little’s Test for Missing Completely At Random (MCAR) (Little, 1988) within Stata Version 15.1 (StataCorp, 2017) to explore the pattern of the missing data in the sleep, psychopathology, and control variables. The results indicated that the missing values within the complete dataset including all variables were not MCAR ($X^2= 121.49$; $df= 25$; $p\leq .001$). In general, a recommended approach is then to perform a multilevel multiple imputation, as it may decrease the risk of bias as compared to listwise deletion if data is either missing at random or not missing at random (van Ginkel et al., 2020).

After the data cleaning steps, we divided our total sample into stratified subsamples, a 10% training (N=56) and a 90% (N=502) testing dataset. Then, we conducted multiple multilevel imputation for the sleep duration and sleep quality variables within both subsamples, as this approach is currently recommended for multilevel data, where measurement days are nested within participants (van Ginkel et al., 2020). To conduct a multiple imputation for multilevel models, we used the ‘mitml’-package (Grund et al., 2019) in R (R Core Team, 2019). Ten imputed datasets were created.

Knot selection for cubic splines with the training subsample (n=56)

To assess the non-linear relationship between sleep duration and psychopathology, the linear regression model included cubic splines. The number of selected knots varies depending on the data, therefore we compared the model fit across three different models using 3, 4, and 5 knots in R (R Core Team, 2019). For each of the 10 imputed datasets, we computed the Bayesian information Criterion (BIC), Akaike information criterion (AIC) values, and the mean squared prediction error (MSPE). The MSPE represents the expected differences between the predicted and observed values of a function which can be used for k-fold cross-validation (Hastie et al., 2008). In this approach, the sample is divided into different subsets (k) which are used to estimate the model for each fold. During this procedure, in each fold the kth subset is kept aside for model validation, allowing the subsets to be cross-validated against each other. Afterwards, the computed MSPE values are averaged across the folds (Hastie et al., 2008). In this study, we used 10-fold cross-validation in each imputed dataset to compute the MSPE value, which was later pooled across all imputed datasets (Hastie et al., 2008). The best fit of the model was determined using a combined approach considering the lowest AIC, BIC, and MSPE value across all imputed datasets (Hagemann et al., 2021).

Conducting a two-step simulation-based power analysis

To the best of our knowledge, no previous study has assessed the associations between objectively measured sleep duration and quality, and psychopathology, within an adolescent sample. Given that no comparable previous literature is available to estimate the distribution of sleep duration and sleep quality and also the estimated effect sizes for the relationships between the sleep variables and psychopathology are unknown, we took a two-step simulation-based approach (Lane & Hennes, 2018) for calculating the power for H1, H2, and H3.

In step 1, within the random 10% training subsample ($n = 56$), two separate linear regressions as outlined below were run in order to get the required effect sizes and model parameters. Second, these model parameters and effect sizes were used as input for 1000 Monte Carlo simulations, which sample size was set to $n = 502$, which is equal to the size of the testing dataset. Of these simulations, the proportion of simulations where the null hypothesis – the effect size of interest being significantly different from zero – was rejected at $p < .05$ was taken as the power. We considered power exceeding .80 to be sufficient.

Estimation of the models in the testing subsample ($n = 502$)

After selecting the number of knots within the training subsample and the imputation of the missing data, we estimated the model to assess our confirmatory hypotheses (1) on the non-linear association between sleep duration and general symptoms of psychopathology, and (2) the non-linear associations of sleep duration in interaction with the age group and symptoms of psychopathology within the 90% testing subsample. Further, we tested our third confirmatory hypothesis on the association between sleep quality and symptoms of psychopathology with a linear regression model. Both linear models were estimated in R (R Core Team, 2019) with the *lm* and *rms* functions using the *rms* package for the cubic splines (Harrell, 2020).

We used Rubin's rules for pooling effect estimates and standard errors across the imputed datasets (Rubin, 1987). The Wald test was calculated using the pooled regression coefficient and standard error. The degrees of freedom and the p-value for the pooled estimate were derived from the adjusted formula version to calculate the df (Van Buuren, 2018) using the MICE package in R (van Buuren & Groothuis-Oudshoorn, 2011).

Results

Characteristics of the training and testing subsample

Table 1 shows the means, standard deviations, and percentages for the variables gender, age, season, sleep duration, sleep quality, BSI-GSI for both the training and the testing subset with and without data imputation.

Table 1. Means, standard deviations, and percentages within the unimputed and imputed training ($n=56$) and testing subsample ($n=502$, across 10 imputed datasets)

	Unimputed training subset ($n=56$)	Unimputed testing subset ($n=502$)	Imputed training subset ($n=56 \times 10$)	Imputed testing subset ($n=502 \times 10$)
Age (years)	13.09 (1.51)	13.13 (1.49)	13.09 (1.51)	13.13 (1.49)
Gender (% female)	64	64	64	64
Season: Winter (%)	43	36	43	36
Season: Spring (%)	45	49	45	49
Season: Autumn (%)	12	15	12	15
Sleep duration (minutes)	472.02 (47.03)	465.85 (54)	470.74 (35.79)	470.51 (36.28)
Sleep quality (0-100)	91.73 (2.16)	92.42 (2.34)	91.79 (1.77)	92.43 (1.93)
BSI-GSI	.83 (.72)	.84 (.62)	.83 (.72)	.84 (.62)
Depressiveness	.84 (.72)	.78 (.60)	.84 (.72)	.78 (.60)
Anxiety	.84 (.72)	.87 (.60)	.84 (.72)	.87 (.60)
Psychoticism	.72 (.72)	.64 (.60)	.72 (.72)	.64 (.60)

Knot selection of cubic splines in training subsample ($n=56$)

We compared the AIC, BIC, and MSPE values from a linear regression model including cubic splines with either, 3, 4, or 5 knots across 10 imputed datasets. Considering all values, the linear regression model including 3 cubic spline knots had the lowest BIC and MSPE values, while there was an equal number of datasets presenting the lowest AIC value in either a linear or a generalized additive model. Standard errors of the MSPE values are reported in **Supplement 6**. Overall, we concluded that the linear model with 3 cubic splines knots provided the best fit (**Table 2** and **Table 3**).

Table 2. *BIC and AIC values for linear models including cubic splines with 3, 4, and 5 knots among ten imputed datasets within imputed training subsample (n=56).*

Dataset	1	2	3	4	5	6	7	8	9	10
BIC										
3 knots	116.56	120.98	122.83	122.13	117.39	124.38	123.01	124.20	124.10	125.00
4 knots	122.92	128.10	130.10	127.19	123.76	126.85	129.04	131.76	126.80	130.66
5 knots	132.50	133.03	133.58	133.94	129.23	130.94	129.90	138.76	134.20	129.6
AIC										
3 knots	133.18	134.75	134.83	142.64	134.80	136.58	135.79	129.11	135.88	129.03
4 knots	136.80	138.31	130.65	145.54	137.08	138.84	136.07	132.35	134.27	127.48
5 knots	140.69	138.24	134.43	147.71	140.18	140.23	138.57	136.17	137.70	128.82

Note: BIC = Bayesian Information Criterion; AIC = Akaike Information Criterion; LM = linear model; LME = linear mixed effects model; GAM = generalized additive model; GAMM general additive mixed effects model

Table 3. *Mean Squared Prediction Error (MSPE) values for a linear including cubic splines with 3, 4, and 5 knots among ten imputed datasets within imputed training subsample (n=56).*

Dataset	1	2	3	4	5	6	7	8	9	10
MSPE										
3 knots	.75	.67	.69	.64	.94	.72	.56	.90	.74	1.08
4 knots	18.65	104.72	63.74	214.68	46.82	50.79	60.00	117.98	36.32	424.25
5 knots	28.14	64.59	43.91	164.32	21.94	237.74	51.52	131.12	31.57	164.37

Note: MSPE = Mean Squared Prediction Error

Simulation-based power analyses

We estimated the power for two hypotheses using a two-step simulation-based power analysis in a sample of $N=502$ using the selected linear model including 3 knots within the cubic regression splines. Our results indicated that a power of $>.99$ would be estimated to reject the Null Hypothesis 1 (“There is no non-linear association between sleep duration and symptoms of psychopathology”). Further, in order to reject the Null Hypothesis 2 (“Complying with age-specific guidelines for sleep duration is associated with symptoms of psychopathology”) a power of $>.095$ was reached. Lastly, for the Null-Hypothesis 3 (“There is no association between sleep quality and symptoms of psychopathology”) the estimated power was $>.99$. Since we consider a power of $.80$ to be sufficient, we conclude that only H1 and H3 are adequately powered, while H2 is underpowered.

Confirmatory analyses

Model fit values for non- linear associations between sleep duration and symptoms of general psychopathology in the testing dataset ($n=502$)

In order to examine the non-linear association between sleep duration and general psychopathology, we conducted a confirmatory analysis using a linear regression model and cubic regression splines. The pooled results across the 10 imputed datasets can be found in **Table 4**. There were no significant associations between sleep duration and general symptoms of psychopathology. To test our hypothesis that complying with sleep guidelines relates to less symptoms of psychopathology, we investigated the interaction effects with a younger (11-14-year-old) and an older (15-17-year-old) age group. There were no significant associations between sleep duration and age group. As expected, we found confounding effects of gender and season. Girls had more symptoms of psychopathology.

Table 4. Associations between sleep duration and psychopathology (GSI-BSI) estimated with a linear regression ($n=502$).

Independent variables	Dependent variable			
	GSI-BSI			
	β (SE)	p	b (SE)	P
Sleep duration	-.00002 (.003)	.99	.0005(.003)	.86
Sleep duration x age group	.002 (.007)	.81	.004 (.007)	.57
Age group	2.22 (2.37)	.14	2.553 (1.68)	.13
Gender	.18 (.06)	.002	.18 (.07)	.006
Season: Spring	-.11 (.06)	.07	-.10 (.06)	.09
Season: Autumn	.04 (.09)	.64	.07 (.10)	.50

Note: in bold = $\alpha < 0.05$; β = unstandardized regression coefficient; b= standardized regression coefficient; SE= standard error.

Model fit values for the associations between sleep quality and general symptoms of psychopathology

In order to assess the association between sleep quality and symptoms of general psychopathology, we conducted a confirmatory analysis using a linear regression model. The pooled results across the 10 imputed datasets can be found in **Table 5**. There were positive, significant associations between sleep quality and general symptoms of psychopathology. As expected, we found confounding effects of gender, age, and season. Fewer symptoms were reported in spring as compared to Winter, while comparing symptoms during Autumn with Winter did not show any significant effects. Since we did not conduct measurements during the Summer due to the students' break, no effects are reported for Summer compared to Winter.

Table 5. Associations between sleep quality and psychopathology (GSI-BSI) estimated with a linear regression in the testing dataset (n=502).

Independent variables	Dependent variable			
	GSI-BSI			
	β (SE)	p	b (SE)	p
Sleep quality ¹	.05 (.02)	.005	.05 (.02)	.02
Age	.06 (.02)	.004	.06 (.02)	.004
Gender	.13 (.06)	.02	.134 (.07)	.05
Season: Spring	-.11 (.06)	.08	-.09 (.06)	.13
Season: Autumn	.01 (.09)	.90	.04 (.10)	.70

Note: in bold = $\alpha < 0.05$; β = unstandardized regression coefficient; SE = standard error

Exploratory analyses

Model fit values for the associations between sleep duration and symptoms of depression, anxiety, and psychoticism in the testing dataset (n=502)

We conducted exploratory analyses using a linear regression model including cubic regression splines. The pooled results across the 10 imputed datasets can be found in **Table 6**. There were no significant non-linear associations between sleep duration with symptoms of depression, anxiety, and psychoticism.

Table 6. Associations between sleep duration and symptoms of depression, anxiety, and psychoticism estimated with a linear regression in the testing dataset (n=502).

Independent variables	Dependent variables					
	Depression		Anxiety		Psychoticism	
	β (SE)	p	β (SE)	P	β (SE)	p
Sleep duration	.0021 (.003)	.63	.001 (.003)	.62	.002 (.003)	.60
Sleep duration x age group	-.00005 (.01)	.79	.005(.007)	.50	-.003 (.008)	.75
Age group	3.48 (2.88)	.22	3.04 (2.17)	.16	.86(2.27)	.71
Gender	.26 (.09)	.003	.15 (.07)	.03	.16 (.07)	.02
Season: Spring	-.12 (.09)	.200	-.14 (.08)	.07	-.14 (.07)	.06
Season: Autumn	.07 (.13)	.61	.04 (.09)	.65	.01 (.10)	.92

Note: in bold = $\alpha < 0.05$; β = unstandardized regression coefficient; SE= standard error.

Model fit values for the associations between sleep quality and symptoms of depression, anxiety, and psychoticism in the testing dataset (n=502)

We conducted exploratory analyses using a linear regression model. The pooled results across the 10 imputed datasets can be found in **Table 7**. There were positive, significant associations between sleep quality and symptoms of depression. However, there were no significant associations between sleep quality and symptoms of anxiety and psychoticism.

Table 7. Associations between sleep quality and symptoms of depression, anxiety, and psychoticism estimated with a linear regression in the testing dataset (n=502).

Independent variables	Dependent variables					
	Depression		Anxiety		Psychoticism	
	β (SE)	p	β (SE)	p	β (SE)	p
Sleep quality	.08 (.03)	.02	.01 (.02)	.52	.04 (.02)	.08
Age	.10 (.03)	<.001	.05 (.02)	.04	.06 (.02)	.003
Gender	.18 (.09)	.06	.15 (.02)	.0455	.12 (.07)	.10
Season: Spring	-.11 (.09)	.24	-.14 (.08)	.08	-1.3 (.08)	.08
Season: Autumn	.02 (.13)	.87	.01 (.10)	.88	-.01 (.10)	.90

Note: in bold = $\alpha < 0.05$; β = unstandardized regression coefficient; SE= standard error.

Discussion

To the best of our knowledge, this is the first study investigating the non-linear relationship between objectively measured sleep duration and symptoms of general psychopathology in a general population sample of adolescents. Our findings showed a non-significant relationship between sleep duration and general symptoms of psychopathology. Therefore, they do not support a non-linear relationship between objectively measured sleep duration and general symptoms of psychopathology. In addition, we explored age-specific effects of sleep duration on psychopathology in groups of both younger (11 to 14-year-old) and older (15 to 17-year-old) adolescents. Both of these research questions are relevant, as evidence on the nature of the relationship between objectively measured sleep duration and psychopathology is inconclusive, and as current sleep guidelines –based on self-report data – are not optimally tailored to prevent maladaptive mental health outcomes (Chaput et al., 2018). Our results showed a non-significant association of sleep duration in interaction with age groups and general psychopathology symptoms. Thus, these findings do not provide evidence for interaction effects between age groups (i.e. younger and older adolescents) and sleep duration in relation to symptoms of psychopathology. As the limitations of self-report data extend to sleep quality as well, we also examined if objectively measured sleep quality relates to fewer symptoms of general psychopathology in the general adolescent population. Our data showed that objectively measured sleep quality was significantly and positively related to symptoms of general psychopathology. These findings imply that adolescents with higher sleep quality experience more symptoms of general psychopathology. These findings are novel, as no previous study has investigated the association of objectively measured sleep quality and general symptoms of psychopathology. Moreover, we also explored both the non-linear relationships of sleep duration and linear associations of sleep quality with symptoms of depression, anxiety, and psychoticism, to facilitate comparability of our results to those from the existing literature.

Sleep duration and general psychopathology.

Our non-significant results are in contrast to previous cross-sectional and longitudinal self-report studies showing a significant association between sleep duration and overall psychopathology symptoms (Hestetun et al., 2018; Matamura et al., 2014). Yet, self-report data may be less reliable, as questionnaire measures are not only generally more prone to overestimating sleep duration in adolescents (Guedes et al., 2016), but are in particular unreliable when measuring sleep in participants experiencing psychopathology symptoms. Matthews and colleagues (2018) showed that the discrepancy between subjective and objective sleep duration measures increased in individuals experiencing depressive symptoms, hostility, and poor health (Matthews et al., 2018). This findings suggests that using objective sleep measures in a general population sample could be more reliable than subjective measures.

In line with our findings, the similarly-sized Trondheim Early Secure Study found that accelerometry-measured short sleep duration at the age of 10 ($n = 557$) was not associated with emotional or behavioural disorders two years later (Ranum et al., 2019). Although our study was cross-sectional, so not directly comparable to the Trondheim Early Secure Study, these similar results underscore that objective findings do not seem to align with findings based on self-report data. As pointed out above, discrepancy between objective and subjective findings may be due to an estimation bias in people with psychopathology symptoms (Matthews et al., 2018). However, it still remains to be seen if this phenomenon also occurs in the general adolescent population.

Overall, the lack of support for our hypothesis adds to an emerging picture from the literature suggesting that objective sleep duration may not relate to psychopathology symptoms in the general adolescent population, as previous self-report studies have assumed (Sadeh et al., 2014). Adolescents sleeping shorter or longer than the current age-specific sleep duration ranges may not necessarily experience adverse effects on mental health, as sleep duration needs vary inter-individually due to genetic and environmental factors (Chaput et al., 2018). On the other hand, it is likely that for most adolescents, the optimal amount of sleep lies within the range of age-appropriate sleep duration recommended by sleep guidelines (Chaput et al., 2018). This implies that factors other than mean

sleep duration may be relevant, such as day-to-day sleep duration variability (i.e. sleep consistency) (Becker et al., 2017). For example, within-person analyses indicate that high variability in objectively measured sleep duration relates to more psychopathology symptoms in adolescents (Bustamante et al., 2020) and subjective well-being than adults (Lemola et al., 2013). Future research is necessary to disentangle the relationship between micro-level changes in sleep and psychopathology symptoms in adolescents – for example by the combined use of accelerometry data and the Experience Sampling Method.

Sleep duration, age, and general psychopathology.

Our results also did not provide evidence for interaction effects between age groups (i.e. younger and older adolescents) and sleep duration in relation to symptoms of general psychopathology. Therefore, these findings do not support the hypothesis that compliance with sleep guidelines relates to fewer symptoms of psychopathology, which may be surprising given that current recommendations from the National Sleep Foundation (Hirshkowitz et al., 2015) relate compliance to maintenance of general health. Again, our findings may speak to the discrepancy between subjective and objective data – as sleep guidelines are based on retrospective (subjective) self-report (Chaput et al., 2018). Moreover, the results of the current study are consistent with a previous statement by Chaput and colleagues (2018) that it is unrealistic to assume that every individual of a given populations needs the same amount of sleep to function well, even though this is what is suggested by sleep guidelines.

Our model validation steps showed that the assumption of homogeneity was violated, indicating that the current non-linear model is not an adequate model to represent the data of our study (**Supplement 3**). To address this violation, we changed the model to a linear regression, which met the assumption of homogeneity in the model validation steps. This model included the variables sleep duration, age, gender, and season, with general psychopathology as the outcome (**Supplement 4**). The findings from this model estimation did not differ from those we derived from the original, non-linear model including interaction effects with age group. Future research may explore which model adequately represents a non-linear relationship between sleep duration, age group and psychopathology within a separate training subsample, before estimating the model in the remainder

of the data. We did not post-register a larger subsample, because we expected the percentage of missing data for the sleep variable to be high. Post hoc, we evaluated that our 10% subsample (n=56) was not large enough to evaluate the model, so larger subsamples (e.g. 50%) may need to be considered.

Further, as our study had insufficient power to investigate the interaction between age, sleep duration, and psychopathology (see results section), the results need to be interpreted cautiously. Given that the current study had a sample size of N=502, yet was still underpowered to test this interaction effect, this suggests that even larger studies are required to robustly test these associations.

Sleep quality and general psychopathology symptoms.

Counterintuitively, our results showed that there was a significant positive association between sleep quality and general symptoms of psychopathology, suggesting that those with better sleep quality experienced more symptoms of psychopathology. These findings are at odds with those from a previous study reporting negative associations between self-reported sleep quality and psychiatric symptoms (Hestetun et al., 2018). It remains challenging to fully interpret these unexpected findings due to several limitations regarding the Fitbit sleep score measure. As such, the Fitbit sleep score comprises number of awakenings, heart rate, and sleep stages. However, these aspects of sleep quality are not recognized by the National Sleep Foundation as objective measures of sleep quality, who define objectively measured sleep quality as including sleep latency, wake after sleep onset, awakenings, and sleep efficiency (Ohayon et al., 2017). Sleep stages, in particular, have not been considered as being representative or a part of sleep quality (Ohayon et al., 2017).

Interpreting the role of sleep stages into the Fitbit sleep score is particularly challenging for many reasons. First, the agreement between sleep stages in Fitbit devices and polysomnography may be low (0.5 – 0.6) in adolescents (Lee et al., 2019), which suggests that sleep stages may not be accurate in the Fitbit score. Second, while it is recognized that for example REM sleep may have a role in symptoms of depression (Palagini et al., 2013), little is known how exactly different aspects of sleep stages (i.e. latency, duration, density) are associated with psychopathology symptoms in adolescents, although more evidence is emerging (Pesonen et al., 2019). Third, it is currently unclear how the Fitbit calculates the sleep score, as the algorithm is not publicly available.

Therefore, the interpretation of our findings needs to be considered in light of the limited conceptualization of the sleep quality construct in our study as compared to agreed objective conceptualizations (Ohayon et al., 2017) and the little agreement of sleep stages with polysomnography.

It could be relevant to further investigate the different dimensions of objective sleep quality, such as sleep efficiency, sleep onset latency, awakenings, and wake after sleep onset in adolescents. Compared to our non-significant findings on sleep duration and psychopathology, the significant findings for sleep quality and general psychopathology may suggest that sleep quality is more relevant for psychopathology symptoms than sleep duration. Previous studies suggest a particularly important role of sleep quality, showing that only sleep quality remained associated with psychopathology when sleep duration was accounted for (Hestetun et al., 2018; Muzni et al., 2021). It has been noted earlier that – at a public health level – sleep quality may be more meaningful to mental health than sleep duration (Bin, 2016). From a practical viewpoint, public health messages to improve sleep quality may be better suited to help individuals than messages related to sleep duration, as they do not address the individual differences in sleep need (Bin, 2016).

Exploratory analyses

Sleep duration, age, and symptoms of depression, anxiety, and psychoticism (exploratory analyses).

Given the limited number of studies exploring the association of objectively measured sleep duration and sleep quality with symptoms of depression, anxiety, and psychoticism, we added a further exploration of these relationships to our study. Similar to the results of general psychopathology, we found that sleep duration was not significantly associated with symptoms of depression, anxiety, and psychoticism. Likewise, sleep duration also did not interact with age groups in our depression, anxiety, and psychoticism analyses. Our findings diverge from large-scale self-report studies, reporting associations between both short and long sleep duration and risk for symptoms of psychoticism (Morishima et al., 2020), depression, and anxiety (Ojio et al., 2016) in the general population of adolescents. As in our confirmatory analyses, these contrasting findings between self-report and objectively-measured could be a function of the differential reliability of subjective and

objective measures (Matthews et al., 2018) or differing measurement periods (Hamann et al., 2019; Slykerman et al., 2019), suggests that future research should explore the effects of different measurement periods in relation to study outcomes.

However, our findings are in line with those from the Auckland Birthweight Collaborative Study (n=871; 50.5% girls), showing that accelerometry-derived sleep duration in 11-year olds was neither cross-sectionally nor longitudinally associated with depressive symptoms at the age of 16 (n=491) (Slykerman et al., 2019). Our findings add to an emerging picture that objectively measured sleep duration is not related to depression symptoms in the general population of adolescents.

Sleep quality, age, and symptoms of depression, anxiety, and psychoticism.

Similar to our findings on general psychopathology, the exploratory analyses showed that objectively measured sleep quality did not relate to symptoms of psychoticism. These findings support those of previous studies in which accelerometry measured sleep quality was not associated with psychotic-like experiences in 11-year-olds (n=814) (Koopman-Verhoeff et al., 2019). Furthermore, we observed a non-significant relationship between objectively measured sleep quality and anxiety symptoms – a relationship that has not previously been investigated in the adolescent general population. Similar to our findings on sleep duration, these findings diverge from those of studies applying subjective measurements (Kaneita et al., 2009).

Counterintuitively, we found that better sleep quality related to more symptoms of depression – a positive association. An additional power analysis indicated that we had sufficient power to investigate this research question ($>.96$). However, the effect size was small ($.06$), suggesting only a weak association. This finding is in contrast to significant negative associations in a small accelerometry study (Hamann et al., 2019) and self-report studies (Kaneita et al., 2009; Raniti et al., 2017). Similar to the findings in general psychopathology, it remains challenging to fully interpret these findings, as we do not know how the sleep quality index was calculated by Fitbit. Yet, understanding these calculations may be important as increased fragmentation of REM-sleep may be associated with more symptoms of depression in adolescents (Pesonen et al., 2019).

Strengths and limitations

Our study has several considerable strengths. First, we used an objective accelerometry measure, as opposed to commonly used self-report measures of sleep. Second, we aimed to ensure maximal rigor by using open science practices (post-registration and open code) to increase the transparency and reproducibility of our work, conducting a simulation-based power analysis, and addressing missing data using multiple imputation methods for multilevel data in order to decrease the risk of bias (van Ginkel et al., 2020). Third, given the absence of a ‘gold standard’ for handling missing Fitbit data (Feehan et al., 2018), our method may be a starting point to help to increase the rigor of future research using Fitbit devices.

Our results also need to be interpreted in light of some limitations. First, our analyses regarding the interaction effects of age group and sleep duration were underpowered. Given that the current study had a sample size of N=502, yet was still underpowered to test this interaction effect, even larger studies are required to robustly test these associations. Second, although we instructed participants to maintain their typical sleep routine during the study, it is possible that some adolescents’ sleep was influenced by responding to early-morning or late-evening experience sampling questionnaires (data not used in the current study). For example, some participants may have got up earlier in the weekend or stayed up later to answer the ESM questionnaire. Third, sleep quality, as assessed via the sleep score in Fitbit, is likely to be less accurate than polysomnography-derived sleep quality measures. In particular, sleep stages, which are included within the composite Fitbit sleep quality index, would benefit from further refinement.

Fourth, the composite Fitbit sleep quality score includes awakenings, heart rate and sleep stages. However, only ‘awakenings’ are currently recognized as an objectively measured aspect of ‘sleep quality’ by the National Sleep Foundation (Ohayon et al., 2017). Future studies may consider conceptualizing sleep quality as per National Sleep Foundation guidelines (Ohayon et al., 2017), which include sleep onset latency, wake after sleep onset, awakening, and sleep efficiency. Fifth, our model validation steps showed that the current non-linear model may not be the optimal model to represent our data, because the assumption of homogeneity was violated (**Supplement 3**). Instead, a linear model may be better (**Supplement 4**). Future research may further explore model fit using a

larger training subsample than in the current study ($n=56$), which seemed too small to assess model assumptions.

Conclusion

To conclude, our study does not provide evidence that objectively measured sleep duration is associated with symptoms of psychopathology in the general adolescent population. However, we found positive, yet small to negligible, associations between sleep quality and depression and general symptoms of psychopathology. Overall, our results suggest that associations between sleep duration and psychopathology symptoms from previous studies may be a function of the subjective, self-report nature of the employed measures, and do not generalize to objectively collected sleep data. To move the field further, future research may explore the different dimensions of sleep quality captured by subjective and objective measures and within-person associations between sleep and psychopathology symptoms.

References

- Becker, S. P., Sidel, C. A., Van Dyk, T. R., Epstein, J. N., & Beebe, D. W. (2017). Intraindividual variability of sleep/wake patterns in relation to child and adolescent functioning: A systematic review. *Sleep Medicine Reviews*, 34, 94–121. <https://doi.org/10.1016/j.smr.2016.07.004>
- Benning, S. D., Bachrach, R. L., Freeman, E., & Wright, A. (2019). The Registration Continuum in Clinical Science: A Guide toward Transparent Practices. *J Abnorm Psychol*, 128(6), 528–540. <https://doi.org/10.1037/abn0000451>
- Beurs, E. D. E., & Zitman, F. G. (2006). De Brief Symptom Inventory (BSI): De betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90. [The Brief Symptom Inventory (BSI): The reliability and validity of a brief alternative of the SCL-90]. *Maandblad Geestelijke Volksgezondheid*, 61, 120–141.
- Bin, Y. S. (2016). Is Sleep Quality More Important than Sleep Duration for Public Health? *Sleep*, 39(9), 1629–1630. <https://doi.org/10.5665/sleep.6078>
- Bustamante, C. M. V., Rodman, A. M., Dennison, M. J., Flournoy, J. C., Mair, P., & McLaughlin, K. A. (2020). Within-person fluctuations in stressful life events, sleep, and anxiety and depression symptoms during adolescence: a multiwave prospective study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 61(10), 1116–1125. <https://doi.org/10.1111/jcpp.13234>
- Cairns, K. E., Yap, M. B. H., Pilkington, P. D., & Jorm, A. F. (2014). Risk and protective factors for depression that adolescents can modify: A systematic review and meta-analysis of longitudinal studies. In *Journal of Affective Disorders* (Vol. 169, pp. 61–75). <https://doi.org/10.1016/j.jad.2014.08.006>
- Chaput, J. P., Dutil, C., & Sampasa-Kanyinga, H. (2018). Sleeping hours: What is the ideal number and how does age impact this? *Nature and Science of Sleep*, 10, 421–430. <https://doi.org/10.2147/NSS.S163071>
- Delaruelle, K., Dierckens, M., Vandendriessche, A., & Deforche, B. (2019). *Studie Jongeren en Gezondheid, Deel 3: gezondheid en welzijn – Uitgelicht: slaap [Factsheet]*. https://www.jongeren-en-gezondheid.ugent.be/wordpress/wp-content/uploads/2020/03/HBSC_201718_Factsheet_slaap.pdf
- Derogatis, L. R. (1993). *BSI Brief Symptoms Inventory: Administration, Scoring, and Procedure Manual (4th Ed.)*. Pearson, NCS.
- Fitbit. (2021a). *How do I track my sleep with my Fitbit device?* https://help.fitbit.com/articles/en_US/Help_article/1314.htm#timeasleep
- Fitbit. (2021b). *What's sleep score in the Fitbit app?* https://help.fitbit.com/articles/en_US/Help_article/2439.htm#what
- Fitzgerald, C. T., Messias, E., & Buysse, D. J. (2011). Teen sleep and suicidality: Results

- from the youth risk behavior surveys of 2007 and 2009. *Journal of Clinical Sleep Medicine*, 7(4), 351–356. <https://doi.org/10.5664/JCSM.1188>
- Franco, P., Putois, B., Guyon, A., Raoux, A., Papadopoulou, M., Guignard-Perret, A., Bat-Pitault, F., Hartley, S., & Plancoulaine, S. (2020). Sleep during development: Sex and gender differences. *Sleep Medicine Reviews*, 51, 101276. <https://doi.org/10.1016/j.smr.2020.101276>
- Galland, B. C., Short, M. A., Terrill, P., Rigney, G., Haszard, J. J., Coussens, S., Foster-Owens, M., & Biggs, S. N. (2018). Establishing normal values for pediatric nighttime sleep measured by actigraphy: a systematic review and meta-analysis. *Sleep*, 41(4). <https://doi.org/10.1093/sleep/zsy017>
- Grund, S., Robitzsch, A., & Luedtke, O. (2019). *mitml: Tools for Multiple Imputation in Multilevel Modeling. R package version 0.3-7*. <https://cran.r-project.org/package=mitml>
- Guedes, L. G., Abreu, G. de A., Rodrigues, D. F., Teixeira, L. R., Luiz, R. R., & Bloch, K. V. (2016). Comparação entre duração do sono autodeclarada e actigrafia entre adolescentes: Diferenças de gênero. *Revista Brasileira de Epidemiologia*, 19(2), 339–347. <https://doi.org/10.1590/1980-5497201600020011>
- Hagemann, N., Kirtley, O. J., Lafit, G., Wampers, M., Achterhof, R., Ermans, K. S. F. M., Hiekkaranta, A. P., Lecei, A., Myin-Germeys, I., & Vancampfort, D. (2021). Objectively measured physical activity and symptoms of psychopathology in general population adolescents from the SIGMA cohort. *Mental Health and Physical Activity*, 21, 100416. <https://doi.org/10.1016/j.mhpa.2021.100416>
- Haghighat, S., Khoshnevis, S., Smolensky, M. H., Diller, K. R., & Castriotta, R. J. (2019). Accuracy of wristband fitbit models in assessing sleep: Systematic review and meta-analysis. In *Journal of Medical Internet Research* (Vol. 21, Issue 11, p. e16273). <https://doi.org/10.2196/16273>
- Hamann, C., Rusterholz, T., Studer, M., Kaess, M., & Tarokh, L. (2019). Association between depressive symptoms and sleep neurophysiology in early adolescence. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 60(12), 1334–1342. <https://doi.org/10.1111/jcpp.13088>
- Harrell, F. E. J. (2020). *rms: Regression Modeling Strategies. R package version 6.1-0*. <https://cran.r-project.org/package=rms>
- Hartmann, J. A., McGorry, P. D., Destree, L., Amminger, G. P., Chasen, A. M., Davey, C. G., Ghie, R., Polari, A., Ratheesh, A., Yuen, H. P., & Nelson, B. (2021). Pluripotential Risk and Clinical Staging: Theoretical Considerations and Preliminary Data From a Transdiagnostic Risk Identification Approach. *Frontiers in Psychiatry*, 11. <https://doi.org/10.3389/fpsy.2020.553578>
- Hastie, T., Tibshirani, R., & Friedman, J. (2008). Model assessment and selection. In *The Elements of Statistical Learning. Data Mining, Inference, and Prediction* (Second).

Springer.

- Hestetun, I., Svendsen, M. V., & Oellingrath, I. M. (2018). Sleep problems and mental health among young Norwegian adolescents. *Nordic Journal of Psychiatry*, 72(8), 578–585. <https://doi.org/10.1080/08039488.2018.1499043>
- Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., Hazen, N., Herman, J., Adams Hillard, P. J., Katz, E. S., Kheirandish-Gozal, L., Neubauer, D. N., O'Donnell, A. E., Ohayon, M., Peever, J., Rawding, R., Sachdeva, R. C., Setters, B., Vitiello, M. V., & Ware, J. C. (2015). National Sleep Foundation's updated sleep duration recommendations: Final report. *Sleep Health*, 1(4), 233–243. <https://doi.org/10.1016/j.sleh.2015.10.004>
- Kaneita, Y., Ohida, T., Osaki, Y., Tanihata, T., Minowa, M., Suzuki, K., Wada, K., Kanda, H., & Hayashi, K. (2007). Association between mental health status and sleep status among adolescents in Japan: A nationwide cross-sectional survey. *Journal of Clinical Psychiatry*, 68(9), 1426–1435. <https://doi.org/10.4088/jcp.v68n0916>
- Kaneita, Y., Yokoyama, E., Harano, S., Tamaki, T., Suzuki, H., Munezawa, T., Nakajima, H., Asai, T., & Ohida, T. (2009). Associations between sleep disturbance and mental health status: A longitudinal study of Japanese junior high school students. *Sleep Medicine*, 10(7), 780–786. <https://doi.org/10.1016/j.sleep.2008.06.014>
- Kessler, R. C., Angermeyer, M., Anthony, J. C., De Graaf, R., Demyttenaere, K., Gasquet, I., Giovanni, De Girolamo, G., Gluzman, S., Gureje, O., Haro, J. M., Kawakami, N., Karam, A., Levinson, D., Mora, M. M. E., Aguilar-Gaxiola, S., Alonso, J., Lee, S., Heeringa, S., ... for the WHO World Mental Health Survey Consortium. (n.d.). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*, 6(3), 168–176.
- Kirtley, O. J., Achterhof, R., Hagemann, N., Hermans, K. S. F. M., Hiekkaranta, A. P., Lecei, A., Boets, B., Henquet, C., Kasanova, Z., Schneider, M., Reiningshaus, U., Viechtbauer, W., & Myin-Germeys, I. (2021). *Initial cohort characteristics and protocol for SIGMA: An accelerated longitudinal study of environmental factors, inter- and intrapersonal processes, and mental health in adolescence*. <https://doi.org/10.31234/osf.io/jp2fk>
- Kline, C. (2013). Sleep Quality. In M. D. Gellman & J. R. Turner (Eds.), *Encyclopedia of Behavioral Medicine* (p. 117). Springer. https://doi.org/https://doi.org/10.1007/978-1-4419-1005-9_849
- Koopman-Verhoeff, M. E., Bolhuis, K., Cecil, C. A. M., Kocavska, D., Hudziak, J. J., Hillegers, M. H. J., Mileva-Seitz, V. R., Reiss, I. K., Duijts, L., Verhulst, F., Luijk, M. P. C. M., & Tiemeier, H. (2019). During day and night: Childhood psychotic experiences and objective and subjective sleep problems. *Schizophrenia Research*, 206, 127–134. <https://doi.org/10.1016/j.schres.2018.12.002>
- Lane, S. P., & Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel

- relationships research. *Journal of Social and Personal Relationships*, 35(1), 7–31.
<https://doi.org/10.1177/0265407517710342>
- Lee, X. K., Chee, N. I. Y. N., Ong, J. L., Teo, T. B., Van Rijn, E., Lo, J. C., & Chee, M. W. L. (2019). Validation of a consumer sleep wearable device with actigraphy and polysomnography in adolescents across sleep opportunity manipulations. *Journal of Clinical Sleep Medicine*, 15(9), 1337–1346. <https://doi.org/10.5664/jcsm.7932>
- Lemola, S., Ledermann, T., & Friedman, E. M. (2013). Variability of Sleep Duration Is Related to Subjective Sleep Quality and Subjective Well-Being: An Actigraphy Study. *PLoS ONE*, 8(8), e71292. <https://doi.org/10.1371/journal.pone.0071292>
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202.
- Liu, B. P., Wang, X. T., Liu, Z. Z., Wang, Z. Y., An, D., Wei, Y. X., Jia, C. X., & Liu, X. (2020). Depressive symptoms are associated with short and long sleep duration: A longitudinal study of Chinese adolescents. *Journal of Affective Disorders*, 263(44), 267–273. <https://doi.org/10.1016/j.jad.2019.11.113>
- Lovato, N., & Gradisar, M. (2014). A meta-analysis and model of the relationship between sleep and depression in adolescents: Recommendations for future research and clinical practice. *Sleep Medicine Reviews*, 18(6), 521–529.
<https://doi.org/10.1016/j.smr.2014.03.006>
- Matamura, M., Tochigi, M., Usami, S., Yonehara, H., Fukushima, M., Nishida, A., Togo, F., & Sasaki, T. (2014). Associations between sleep habits and mental health status and suicidality in a longitudinal survey of monozygotic twin adolescents. *Journal of Sleep Research*, 23(3), 290–294. <https://doi.org/10.1111/jsr.12127>
- Matricciani, L., Paquet, C., Galland, B., Short, M., & Olds, T. (2019). Children’s sleep and health: A meta-review. In *Sleep Medicine Reviews* (Vol. 46, pp. 136–150). W.B. Saunders Ltd. <https://doi.org/10.1016/j.smr.2019.04.011>
- Matthews, K. A., Patel, S. R., Pantescio, E. J., Buysse, D. J., Kamarck, T. W., Lee, L., & Hall, M. H. (2018). Similarities and differences in estimates of sleep duration by polysomnography, actigraphy, diary, and self-reported habitual sleep in a community sample. *Sleep Health*, 4(1), 96–103. <https://doi.org/10.1016/j.sleh.2017.10.011>
- Menghini, L., Yuksel, D., Goldstone, A., Baker, F. C., & de Zambotti, M. (2021). Performance of Fitbit Charge 3 against polysomnography in measuring sleep in adolescent boys and girls. *Chronobiology International*.
<https://doi.org/10.1080/07420528.2021.1903481>
- Michaud, I., & Chaput, J. P. (2016). Are Canadian children and adolescents sleep deprived? *Public Health*, 141, 126–129. <https://doi.org/10.1016/j.puhe.2016.09.009>
- Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J.,

- Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Medicine*, 47(9), 1821–1845. <https://doi.org/10.1007/s40279-017-0716-0>
- Morishima, R., Yamasaki, S., Ando, S., Shimodera, S., Ojio, Y., Okazaki, Y., Kasai, K., Sasaki, T., & Nishida, A. (2020). Long and short sleep duration and psychotic symptoms in adolescents: Findings from a cross-sectional survey of 15 786 Japanese students. *Psychiatry Research*, 293, 113440. <https://doi.org/10.1016/j.psychres.2020.113440>
- Muzni, K., Groeger, J. A., Dijk, D. J., & Lazar, A. S. (2021). Self-reported sleep quality is more closely associated with mental and physical health than chronotype and sleep duration in young adults: A multi-instrument analysis. *Journal of Sleep Research*, 30(1). <https://doi.org/10.1111/jsr.13152>
- Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., & Reininghaus, U. (2018). Experience sampling methodology in mental health research: new insights and technical developments. *World Psychiatry*, 17(2), 123–132. <https://doi.org/10.1002/wps.20513>
- Ohayon, M., Wickwire, E. M., Hirshkowitz, M., Albert, S. M., Avidan, A., Daly, F. J., Dauvilliers, Y., Ferri, R., Fung, C., Gozal, D., Hazen, N., Krystal, A., Lichstein, K., Mallampalli, M., Plazzi, G., Rawding, R., Scheer, F. A., Somers, V., & Vitiello, M. V. (2017). National Sleep Foundation’s sleep quality recommendations: first report. *Sleep Health*, 3(1), 6–19. <https://doi.org/10.1016/j.sleh.2016.11.006>
- Ojio, Y., Nishida, A., Shimodera, S., Togo, F., & Sasaki, T. (2016). Sleep duration associated with the lowest risk of depression/anxiety in adolescents. *Sleep*, 39(8), 1555–1562. <https://doi.org/10.5665/sleep.6020>
- Palagini, L., Baglioni, C., Ciapparelli, A., Gemignani, A., & Riemann, D. (2013). REM sleep dysregulation in depression: State of the art. In *Sleep Medicine Reviews* (Vol. 17, Issue 5, pp. 377–390). W.B. Saunders. <https://doi.org/10.1016/j.smr.2012.11.001>
- Paruthi, S., Brooks, L. J., D’Ambrosio, C., Hall, W. A., Kotagal, S., Lloyd, R. M., Malow, B. A., Maski, K., Nichols, C., Quan, S. F., Rosen, C. L., Troester, M. M., & Wise, M. S. (2016). Recommended amount of sleep for pediatric populations: A consensus statement of the American Academy of Sleep Medicine. *Journal of Clinical Sleep Medicine*, 12(6), 785–786. <https://doi.org/10.5664/jcsm.5866>
- Patton, G. C., Coffey, C., Romaniuk, H., Mackinnon, A., Carlin, J. B., Degenhardt, L., Olsson, C. A., & Moran, P. (2014). The prognosis of common mental disorders in adolescents: A 14-year prospective cohort study. *The Lancet*, 383(9926), 1404–1411. [https://doi.org/10.1016/S0140-6736\(13\)62116-9](https://doi.org/10.1016/S0140-6736(13)62116-9)
- Pesonen, A. K., Gradisar, M., Kuula, L., Short, M., Merikanto, I., Tark, R., Räikkönen, K., & Lahti, J. (2019). REM sleep fragmentation associated with depressive symptoms and

- genetic risk for depression in a community-based sample of adolescents. *Journal of Affective Disorders*, 245, 757–763. <https://doi.org/10.1016/j.jad.2018.11.077>
- Quante, M., Wang, R., Weng, J., Kaplan, E. R., Rueschman, M., Taveras, E. M., Rifas-Shiman, S. L., Gillman, M. W., & Redline, S. (2019). Seasonal and weather variation of sleep and physical activity in 12–14-year-old children. *Behavioral Sleep Medicine*, 17(4), 398–410. <https://doi.org/10.1080/15402002.2017.1376206>
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org>
- Raniti, M. B., Allen, N. B., Schwartz, O., Waloszek, J. M., Byrne, M. L., Woods, M. J., Bei, B., Nicholas, C. L., & Trinder, J. (2017). Sleep Duration and Sleep Quality: Associations With Depressive Symptoms Across Adolescence. *Behavioral Sleep Medicine*, 15(3), 198–215. <https://doi.org/10.1080/15402002.2015.1120198>
- Ranum, B. M., Wichstrøm, L., Pallesen, S., Falch-Madsen, J., Halse, M., & Steinsbekk, S. (2019). Association Between Objectively Measured Sleep Duration and Symptoms of Psychiatric Disorders in Middle Childhood. *JAMA Network Open*, 2(12), e1918281. <https://doi.org/10.1001/jamanetworkopen.2019.18281>
- Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys*. Wiley.
- Sadeh, A., Tikotzky, L., & Kahn, M. (2014). Sleep in infancy and childhood: Implications for emotional and behavioral difficulties in adolescence and beyond. *Current Opinion in Psychiatry*, 27(6), 453–459. <https://doi.org/10.1097/YCO.0000000000000109>
- Slykerman, R. F., Thompson, J. M. D., Coomarasamy, C., Wall, C. R., Waldie, K. E., Murphy, R., & Mitchell, E. A. (2019). Early adolescent physical activity, sleep and symptoms of depression at 16 years of age. *Acta Paediatrica, International Journal of Paediatrics*, August, 1–6. <https://doi.org/10.1111/apa.15140>
- StataCorp. (2017). *Stata Statistical Software: Release 2015*. StataCorp LLC.
- Van Buuren, S. (2018). *Flexible Imputation of Missing Data* (Second). Chapman & Hall/CRC.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3), 1–67. <https://www.jstatsoft.org/v45/i03/>
- van Ginkel, J. R., Linting, M., Rippe, R. C. A., & van der Voort, A. (2020). Rebutting Existing Misconceptions About Multiple Imputation as a Method for Handling Missing Data. *Journal of Personality Assessment*, 102(3), 297–308. <https://doi.org/10.1080/00223891.2018.1530680>

Appendix

Supplement 1

Inclusion and exclusion criteria

Our dataset consisted of daily summaries of the variables sleep duration and sleep quality. Common proxies of non-wear time at night (e.g. heart rate, accelerometer count) were not available.

Alternatively, we decided that when both sleep duration and sleep quality were simultaneously recorded as 'zero' within the same measurement day, they were labelled as missing. These missing values may have occurred either due to non-wear of the device at night, or due to syncing too late. In that case, the existing data gets erased to provide storage for collecting more recent data. The device cannot store data for longer than 7-9 days and if that period is exceeded, the data collected on day 1 and 2 (sleep and active minutes) will be converted to zero values, in order to store data on the measurement days 9, 10, 11, and so on. This process was confirmed by a Fitbit customer service assistant.

Data exclusion

From our in-lab pilot testing, we also concluded that if the device was briefly removed during the day for at least one hour (due to showering or other reasons) then replaced, *but not worn during the night of the same day*, the device would incorrectly log this period of non-wear time as 'sleep' time. To avoid that short periods of 'sleep' occurred in our data, we excluded these brief 'sleep' periods using the following criterion: Any sleep time that was recorded before 9pm and with a duration of 200 minutes. Additionally, this decision was made since we wanted to assess the effects of nighttime sleep.

Also, if during all 8 measurement days, sleep duration and sleep quality were recorded as zero, all data from this participant was excluded from the analysis. Lastly, as this study investigates sleep in 11-17-year-olds, any participants older than 17 (N=26) were excluded from the analysis.

Supplement 2

Deviations from the post-registration

1. Change from gam model to linear model

As we aimed to impute the missing values of sleep duration, we were not able to estimate generalized additive models using the gam function of the mgcv package, as this package does not support the use of multilevel multiple imputed datasets. Thus, we changed the model to a mixed effects linear model using cubic splines to account for the non-linear effects. As the ICCs of the linear mixed effects model were zero, we simplified the model and did not take the different clusters for school into account anymore.

2. Additional, exploratory analyses for the subscales depression, anxiety, and psychoticism

Since previous studies have mainly investigated the effects of sleep duration and quality on depression, anxiety, and psychoticism, we conducted exploratory analyse for these subscales.

3. Model validation

In our post-registration, we stated that we would check models met the assumption of independence. However, it was not possible for us to pool the imputed results within a plot, so we dropped this and evaluated the assumption of homogeneity only.

4. Imputation of the BSI data

We had stated in our post-registration that imputation would not be feasible, since 60% of the sleep data was missing. However, after we had excluded participants with less than 4 valid days, the 32% missingness of sleep duration and sleep quality seemed a justifiable amount of missingness to impute. Furthermore, based upon further reading after our post-registration, we decided that imputation of the BSI-GSI and BSI subscale variables depression, anxiety, and psychoticism is in general a better and more robust approach than listwise deletion (van Ginkel et al., 2020). We report the findings of the non-imputed BSI and Fitbit variables in the **Supplement 5**.

5. Additional model selection methods

In our post-registration, we stated that if AIC and BIC contradicted each other, we would opt for the BIC values. However, upon further reading, we realised that this approach was incorrect. As the

results from the AIC and BIC estimation did not favour a model, we also used the MSPE value calculated via k-fold cross-validation for additional model selection information.

Supplement 3

Model validation

In line with our post-registration, we checked for homogeneity by plotting the fitted values against the residuals in both models. In the first model, estimating the non- linear associations between sleep duration and symptoms of general psychopathology, we see that the plot shows clear patterns of the residuals. We therefore concluded that the assumption of homogeneity was violated and heteroscedasticity existed (see **Fig.1**). Similarly, we checked our second model estimating the linear associations between sleep quality and symptoms of general psychopathology (see **Fig.2**). Based on the lack of a pattern in the spread, we concluded that the assumption of homogeneity was confirmed.

Fig. 1. *Fitted values against residuals for non- linear associations between sleep duration and symptoms of general psychopathology in the testing dataset (n=502)*

Fig. 2. *Fitted values against residuals for linear associations between sleep quality and symptoms of general psychopathology in the testing dataset (n=502)*

Fig. 2. Fitted values against residuals for linear associations between sleep quality and symptoms of general psychopathology in the testing dataset ($n=502$)

Supplement 4

Since the assumption of homogeneity was violated for the model estimating the non-linear association between sleep duration and psychopathology, we conducted several exploratory analyses in search of an alternative model which demonstrated homogeneity of residual spread. We took several steps to simplify the post-registered model, which included stepwise removal of (in order) (1) cubic splines for the variable sleep duration, (2) interaction effects for sleep duration and age group, and the (3) variable age group,. The model kept showing a clear pattern of residuals until we reached the simplest model: sleep duration, age (as a continuous variable), gender, and season predicting psychopathology. In this model, the residuals indicated an equal spread (**Fig. 3**).

Fig. 3. *Fitted values against residuals for the linear associations between duration and symptoms of general psychopathology in the testing dataset (n=502)*

Supplement 5

For transparency, we report the results of our original post-registered model here in which we did not impute BSI variables. The results from a model estimating the relationship between sleep quality and psychopathology on a dataset in which BSI variables were not imputed were different from our main findings using imputed BSI data. The imputed model showed a significant relationship between sleep quality and psychopathology, while the non-imputed model did not (**Table 8**). In order to evaluate if the imputation of the BSI variable was successful, we compared the distribution of the imputed variable compared to the non-imputed variable. The results indicated that the imputation model was correct. Overall, the fact that the observed dataset is not MCAR (see Methods section), could explain the differences in results between listwise deletion in the observed dataset and multiple imputation. As van Ginkel state (2020), it could imply that results obtained from listwise deletion in the observed dataset are less trustworthy, whilst the results from multiple imputation are not. The authors argue that even if the data is NMAR, they may produce less biased results than in listwise deletion (van Ginkel et al., 2020). However, further empirical research needs to be conducted to understand if listwise deletion or multiple imputation of the BSI-GSI variable is more appropriate in this dataset.

Table 7. Associations between sleep duration and psychopathology (GSI-BSI) estimated with a linear regression in non-imputed testing dataset (n=502).

Independent variables	Dependent variable			
	GSI-BSI			
	β (SE)	p	b (SE)	p
Sleep duration	-.00002 (.003)	.99	.002 (.003)	.48
Sleep duration x age group	.006 (.006)	.47	.005 (.008)	.46
Age group	3.53 (2.37)	.14	3.43 (2.30)	.14
Gender	.21 (.06)	.001	.12 (.07)	.001
Season: Spring	-.15 (.07)	.03	-.15 (.07)	.03
Season: Autumn	.04 (.09)	.71	.04 (.09)	.71

Note: in bold = $\alpha < 0.05$; β = unstandardized regression coefficient; SE= standard error

Table 8. Associations between sleep quality and symptoms of depression, anxiety, and psychoticism estimated with a linear regression in non-imputed testing dataset (n=502).

Independent variables	Dependent variable			
	GSI-BSI			
	β (SE)	p	b (SE)	p
Sleep quality	.03 (.02)	.06	.06 (.04)	.09
Age	.07 (.02)	.0002	.28 (.08)	.0005
Gender	.18 (.07)	.004	.10 (.03)	.003
Season: Spring	-.15 (.07)	.03	-.14 (.07)	.04
Season: Autumn	-.009 (.09)	.92	.03 (.09)	.72

Note: in bold = $\alpha < 0.05$; β = unstandardized regression coefficient; SE= standard error

Table 9. Associations between sleep duration and symptoms of depression, anxiety, and psychoticism estimated with a linear regression in non-imputed testing dataset (n=502).

Independent variables	Dependent variables					
	Depression		Anxiety		Psychoticism	
	β (SE)	p	β (SE)	P	β (SE)	p
Sleep duration	.0001 (.004)	.97	.0009 (.003)	.79	.001 (.003)	.71
Sleep duration x age group	.002 (.01)	.79	.007 (.01)	.48	-.002 (.08)	.75
Age group	5.04 (2.88)	.08	3.82 (2.88)	.20	1.02 (2.69)	.71
Gender	.30 (.09)	<.001	.19 (.07)	.006	.19 (.07)	.006
Season: Spring	-.15 (.09)	.10	-.17 (.08)	.02	-.17 (.08)	.02
Season: Autumn	.08 (.13)	.52	.05 (.10)	.62	.007 (.11)	.94

Note: in bold = $\alpha < 0.05$; β = unstandardized regression coefficient; SE= standard error

Table 10. Associations between sleep quality and symptoms of depression, anxiety, and psychoticism estimated with a linear regression in the testing dataset (n=502).

Independent variables	Dependent variables					
	Depression		Anxiety		Psychoticism	
	β (SE)	p	β (SE)	p	β (SE)	p
Sleep quality	.06 (.02)	.03	.007 (.02)	.70	.03 (.02)	.10
Age	.13 (.03)	<.001	.05 (.02)	.02	.07 (.02)	.001
Gender	.24 (.09)	.006	.20 (.07)	.005	.16 (.07)	.02
Season: Spring	-.16 (.09)	.09	-.17 (.08)	.03	-1.8 (.07)	.02
Season: Autumn	.006 (.13)	.96	.02 (.10)	.90	-.03 (.10)	.77

Note: in bold = $\alpha < 0.05$; β = unstandardized regression coefficient; SE= standard error

Supplement 6

Table 3. *Standard error values of the Mean Squared Prediction Error values for linear models including cubic splines with 3, 4, and 5 knots among ten imputed datasets within imputed training subsample (n=56).*

Dataset	1	2	3	4	5	6	7	8	9	10
MSPE SE										
3 knots	.36	.26	.30	.31	.52	.31	.27	.40	.37	.72
4 knots	23.43	124.81	61.86	412.53	66.01	80.78	90.26	197.07	67.89	833.22
5 knots	49.65	92.44	67.00	268.04	29.38	464.13	90.57	175.61	44.05	397.89

Note: MSPE = Mean Squared Prediction Error; SE =standard errors