


RESEARCH

Open Access



# The sleep loss insult of Spring Daylight Savings in the US is observable in Twitter activity

Kelsey Linnell<sup>1\*</sup> , Michael Arnold<sup>1</sup>, Thayer Alshaabi<sup>1</sup>, Thomas McAndrew<sup>2</sup>, Jeanie Lim<sup>3</sup>, Peter Sheridan Dodds<sup>1</sup> and Christopher M. Danforth<sup>1\*</sup>

\*Correspondence:  
klinnell@uvm.edu; chris.  
danforth@uvm.edu

<sup>1</sup> Computational Story Lab,  
Vermont Complex Systems  
Center, MassMutual Center  
of Excellence for Complex  
Systems & Data Science,  
University of Vermont,  
Burlington, USA  
Full list of author information  
is available at the end of the  
article

## Abstract

Sleep loss has been linked to heart disease, diabetes, cancer, and an increase in accidents, all of which are among the leading causes of death in the United States. Population-scale sleep studies have the potential to advance public health by helping to identify at-risk populations, changes in collective sleep patterns, and to inform policy change. Prior research suggests other kinds of health indicators such as depression and obesity can be estimated using social media activity. However, the inability to effectively measure collective sleep with publicly available data has limited large-scale academic studies. Here, we investigate the passive estimation of sleep loss through a proxy analysis of Twitter activity profiles. We use “Spring Forward” events, which occur at the beginning of Daylight Savings Time in the United States, as a natural experimental condition to estimate spatial differences in sleep loss across the United States. On average, peak Twitter activity occurs 15 to 30 min later on the Sunday following Spring Forward. By Monday morning however, activity curves are realigned with the week before, suggesting that the window of sleep opportunity is compressed in Twitter data, revealing Spring Forward behavioral change.

**Keywords:** Sleep, Sociotechnical systems, Daylight Savings, Twitter, Public health, Twitter behavioral pattern

## Introduction

The American Academy of Sleep Medicine recommends adults sleep 7 or more hours per night [1]. However, studies show only 2/3 of adults sleep for this length of time consistently. In 2014, the Centers for Disease Control and Prevention’s (CDC’s) Behavioral Risk Factor Surveillance System suggested that between 28% and 44% of the adult population of each state received less than the recommended 7 h of sleep [2]. Despite the scientific consensus that adequate sleep is essential to health, many adults are sleeping less than 7 h a night on average—a state referred to as *short sleep*. Results from the most recent National Health Interview Survey determined that since 1985, the age-adjusted average sleep duration has decreased, and the percentage of adults who experience short sleep, on average, rose by 31% [3].

Because adequate sleep is necessary for optimal cognition, short sleep is adverse to productivity and learning, and reduces the human capacity to make effort-related choices such as whether to take precautionary safety measures [4–6]. Short sleep's impact on human cognition is harmful in the workplace, and poses a pronounced and distinct threat to public safety when operating a vehicle [7–10]. Short sleep is linked to increased risk of serious health conditions, including heart disease, obesity, diabetes, arthritis, depression, strokes, hypertension, and cancer [2, 11, 12], and a recent study found that disrupted sleep is also associated with DNA damage [13]. The link between sleep loss and cancer is so strong that the World Health Organization has classified night shift work as “probably carcinogenic to humans” [14]. Socio-economic status is positively correlated with quality of sleep [15–18]. Due to such detrimental effects, and high prevalence among the population, insufficient sleep accounts for between \$280 and over \$400 billion lost in the United States every year [19].

Accurately measuring short sleep in a large population is difficult, and there is often a trade-off between accuracy and the size of the study. Polysomnography—considered the most accurate way to measure sleep—can only measure an individual's sleep patterns in a controlled laboratory setting [20, 21]. Large studies have relied on participants recording their own sleep, but suffer from reporting bias [2, 22, 23].

Wearable technology can measure short sleep at the population scale, and has the potential to measure short sleep accurately enough to study its association with adverse health risks [4, 20, 24, 25]. One recent large sleep study enrolled 31,000 participants and used sleep data from wearable devices along with participant's interactions with a web based search engine to compare sleep loss and performance [4]. The authors [4] showed that measurements of cognitive performance (including keystroke and click latency) vary over time, follow a circadian rhythm, and are related to the duration of participant's sleep, results that closely mirrored those from laboratory settings and validated their methodology. Another study using wearables was able to analyze nine metrics of sleep, including social jetlag, duration, and variability for 69,650 individuals [25]. The authors' analysis of these metrics found gendered differences in sleep behaviors across the cohort [25].

While promising in the long run, present studies that use wearable devices have limitations. To infer from wearables that individuals are sleeping, data must first go through a pipeline of preprocessing, feature extraction and classification. The pipeline for processing sleep data is typically proprietary and dependent on the specific wearable used, and changes to how data is processed can impact results [26]. Moreover, validation studies have yet to explore the effectiveness of these devices across genders, ages, culture, and health [26].

Social media may be an alternative way to measure sleep disturbances in a large population, for example by studying the link between screen time and sleep [27, 28]. Researchers have found that Tweeting behavior can reveal “sleep-wake” behavior for individuals as well as cities [29, 30]. In particular, the correlation between sustained low activity on Twitter and sleep time as measured by conventional surveys has been validated against data collected from the CDC on sleep deprivation [27]. The relationship between time of onset of Twitter activity and wake time has been used to explore and demonstrate social jetlag—the discrepancy between weekend and weekday sleep behavior [27, 31]. Other

work has shown evidence of an increase in a user's smart phone screen time as being associated with an increase in short sleep [28]. Other mental and physical characteristics have been measured from sociotechnical systems. Several instruments developed by members of our research group including the Hedonometer [32], which measures population sentiment through tweets, and the Lexicocalorimeter [33], which measures caloric balance at the state level, have demonstrated an ability to infer population-scale health metrics from Twitter data. Circadian rhythms in mood and cognitive processes have also been inferred from tweets [34, 35]. Twitter data has also been used to identify users who experience sleep deprivation and study the ways their social media interactions differ from others [36].

In urban, industrialized societies where social timing is synced to clock time, Daylight Savings- a biannual sudden upset to clock time- creates behavioral stability across seasons [37, 38]. The onset of DST, Spring Forward, is associated with a 1 h sleep disruption due to the disconnect between the “human clock” and the mechanical clock [39]. Past work has used Daylight Savings as a natural experiment to show that a 1 h collective sleep loss event has large and quantifiable effects on health, safety, and the economy [40–43], with two striking findings being a 1 day increase in heart attacks by 24% and a loss of \$31 billion on the NYSE, AMEX, and NASDAQ exchanges in the United States [40, 44].

We hypothesize here that sleep loss is measurable in behavioral patterns on Twitter, and changes in population-scale sleep patterns due to Spring Forward can be observed through changes in these behavioral patterns. In what follows we describe the process by which we used the local time of tweet posting to explore patterns in posting frequency relative to time of day, and how these patterns were affected by the clock shift known as Spring Forward. The data is described in detail, followed by the specific methodologies employed to analyze the patterns in the frequency of posting. Then, we visualize and describe the results before concluding with a discussion of limitation and implications.

## Materials and methods

### Data

We collected a 10% random sample of all public tweets—offered by Twitter's Decahose API—for Sundays and Mondays in the 4 weeks leading up to, the week of, and the 4 weeks following Spring Forward events during the years 2011–2014. Spring Forward is defined as the instantaneous clock adjustment from 2 a.m. to 3 a.m. on the second Sunday of March each year. We included tweets in the study if the user who created the tweet reported living in the U.S. in their bio, or if the tweet was geo-tagged to a GPS coordinate within the U.S. [45]. With these conditions, we ended up selecting approximately 7% of the messages in the Decahose random sample for analysis [46]. The sample was composed of 13.1 million tweets.

Twitter provided the time-zone from which each message was posted during the period from 2011 to 2014 (for privacy purposes, Twitter discontinued publication of time zone information in 2015). We used the time-zone to determine the local time of posting for each tweet. Tweets for which the time-zone was incompatible with the assigned location were discarded. This process enabled us to analyze the method on data

for which the local time of posting is known. We binned tweets by 15 min increments according to the local time of day they were posted.

### Experimental setup

The Spring Forward event of Daylight Savings was used as a natural experiment in which the control is behavior prior to the event, and the experiment is behavior directly after the clock change and known sleep loss event. Change in Twitter posting behavior was observed in this experiment. To estimate behavioral change associated with Daylight Savings, we partitioned tweets into various groups, primarily a “Before Spring Forward” (BSF) group and a “Spring Forward” (SF) group. To establish a convenient ‘control’ pattern of behavior, all tweets posted on any of the four Sundays before the Spring Forward event were classified as “Before Spring Forward” tweets. We classified the ‘experimental’ set of tweets posted on the Sunday coincident with the Spring Forward event as “Spring Forward”. The above classification created, for every year, a 4:1 matching of before to week of Spring Forward activity. We analyzed tweets posted 1–4 weeks following Spring Forward separately to quantify relaxation to the original behavior.

### Analysis

We binned tweets by time in 15 min intervals starting at the top of the hour, and normalized their frequencies by dividing by the total number of tweets posted on the corresponding day. In this way, we establish a discrete description of the posting volume over the course of a typical 24-h period.

We averaged the Before Spring Forward tweets over the four Sundays, and the 4 years as follows:

$$T_{\text{BSF}}(k) = (4 \times 4)^{-1} \sum_{Y=2011}^{2014} \sum_{S=1}^4 \frac{C_{YS}(k)}{C_{YS}},$$

where  $C_{YS}(k)$  is the number of tweets in the  $k^{\text{th}}$  15 min interval of the  $S^{\text{th}}$  Sunday of year  $Y$ ,  $C_{YS}$  is the total number of tweets posted on that Sunday and year, and  $T_{\text{BSF}}(k)$  is the average fraction of tweets posted in the  $k^{\text{th}}$  15 min interval of a Sunday prior to Spring Forward,

We also normalized the Spring Forward tweets against daily activity:

$$T_{\text{SF}}(k) = (4)^{-1} \sum_{Y=2011}^{2014} \frac{C_Y(k)}{C_Y}.$$

These averages enabled us to aggregate more data, building a more reliable pattern of daily activity, and decrease the susceptibility to daily variation. To reduce noise that could depend on our choice of bin size and spatial scale, we smoothed normalized tweet activity using Gaussian Process Regression (GPR) [47, 48]. We fit a GPR with a squared exponential kernel and characteristic length scale of 150 min (a total of 10 bins of size 15-min) to normalized tweets. We chose a characteristic length of 150 min for consistency with previous work [27]. Tikhonov regularization with an  $\alpha$  penalty tuned manually to 0.1 was included when finding weights  $\omega_k$  to prevent overfitting [48]. GPR yielded a smooth behavioral curve,  $B(t)$ , of the functional form:

$$B(t) = \sum_{k=1}^{96} \omega_k \exp \left[ -\frac{1}{2} k \left( \frac{t}{150}, \frac{t_k}{150} \right)^2 \right],$$

where  $\omega_k$  is a weight determined by the regression process,  $k$  is the squared-exponential kernel (commonly called a radial basis),  $t$  is the time in minutes since midnight (00:00), and  $t_k$  is the  $k^{th}$  15 min interval of the day, i.e.  $t_5$  corresponds to 75 min past midnight, or 1:15 a.m. The sum to 96 refers to the number of 15 min intervals in a single 24 h period.

We generated behavioral curves  $B(t)$  for the BSF and SF groups by state, and for the U.S. in aggregate. To estimate behavioral change induced by a Spring Forward event, we calculate two quantities from the behavioral curves: (i) the time of peak activity and (ii) the time of the inflection point between the peak and trough. The inflection point is referred to as a ‘twinflection’ point, and represents a point of diminishing losses in Twitter activity for the night. Peak shift is defined as:

$$\arg \max_t \{B_{SF}(t)\} - \arg \max_t \{B_{BSF}(t)\}$$

and twinflexion shift is defined as:

$$\arg \min_{t \in N} \{B'_{SF}(t)\} - \arg \min_{t \in N} \{B'_{BSF}(t)\},$$

where  $N = \{t : \arg \max_t B(t) < t < \arg \min_t B(t)\}$ . We were able to reliably measure peak activity and twinflexion because behavioral curves exhibited a consistent diurnal wave structure: a rise in the evening corresponding to peak Twitter posting activity, followed by a trough during typical sleeping hours, and a plateau throughout the day. Contraction of the trough associated with sleeping hours is considered to be reflective of lost sleep opportunity, and may indicate sleep loss itself.

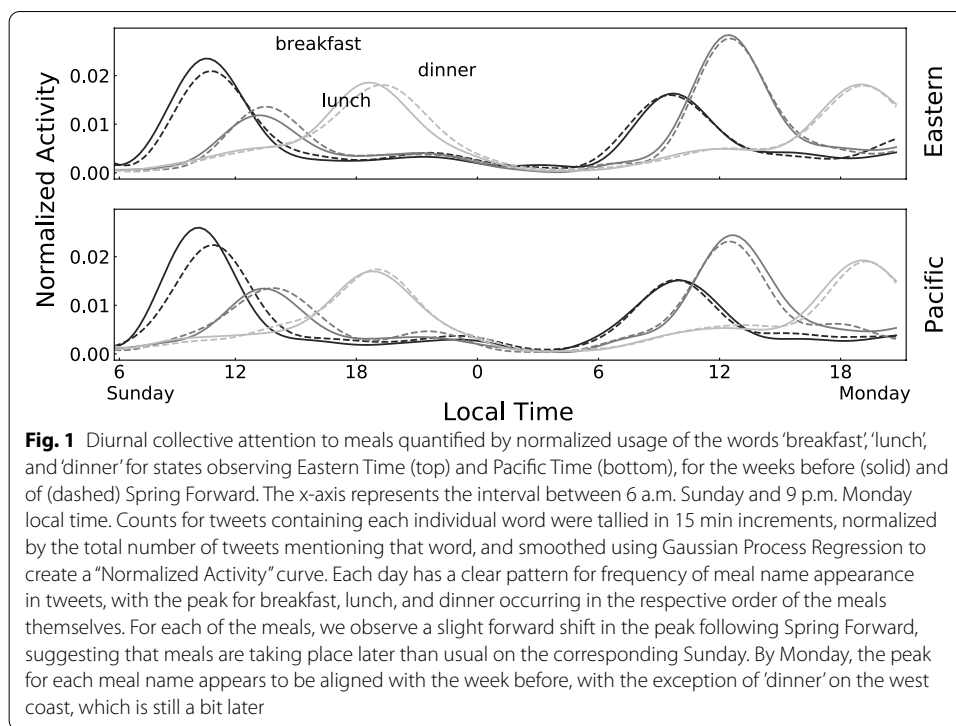
We measured the loss of sleep opportunity by calculating the peak and twinflexion times for the 4 weeks Before Spring Forward and the week of Spring Forward itself. We then characterize differences between the BSF and SF measures for each state, and for the total U.S., as a proxy for sleep loss.

## Results

Our overall finding is that peak Twitter activity occurs 15–30 min later on the Sunday evening immediately following Spring Forward for most states, with this shift varying among states. By Monday morning, activity is back to normal, suggesting that the window of sleep opportunity is visibly compressed in Twitter behavior.

In Fig. 1, we plot  $B(t)$  for the subset of posts containing the words ‘breakfast’, ‘lunch’, and ‘dinner’ for the period beginning 6 a.m. on Sunday and ending 9 p.m. on Monday, both before (solid) and the weeks of (dashed) Spring Forward events. These curves were constructed for states observing Eastern Time (top row) and Pacific Time (bottom row). These regions were chosen as they are the zones with the greatest spatial difference among zones with significant data density. Observing a shift in behavior for each assures us that these shifts are not limited to a particular geographic region of the country.

Meal-related language reveals a daily pattern of behavior in which peak volume occurs around the time that meal typically takes place. On an average Sunday, breakfast is most



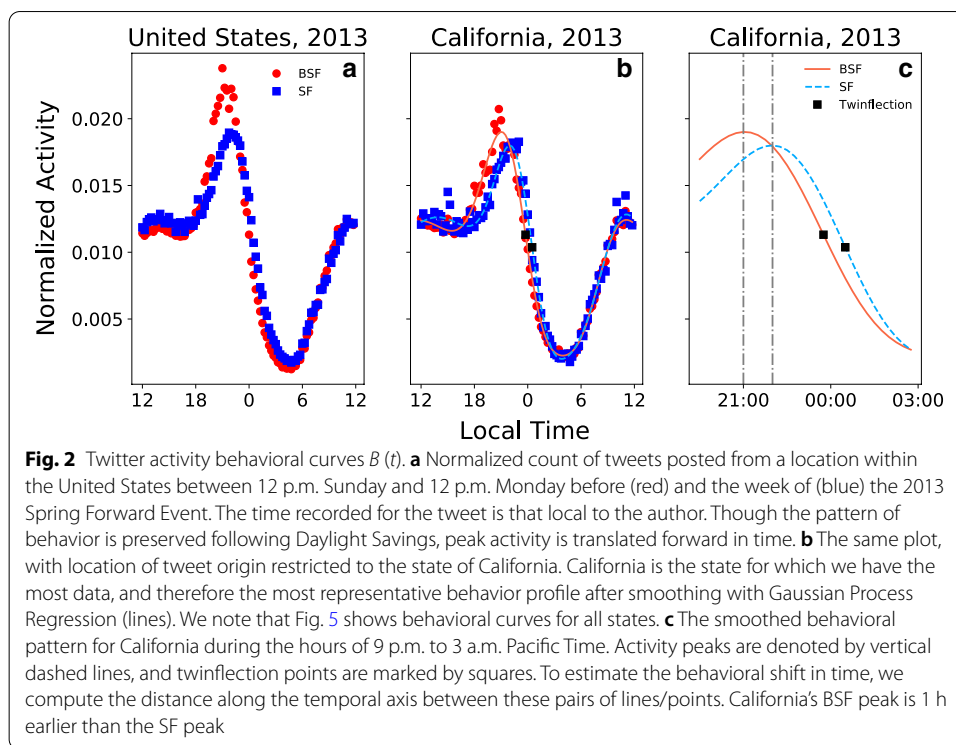
mentioned at 10:30 a.m., lunch at 1:15 p.m., and dinner at 6:45 p.m. in Eastern Time Zone states (see Fig. 1). On the average Monday, breakfast mentions peak at 10:45 a.m., lunch peaks at 1:30 p.m., and dinner at 7:15 p.m. Breakfast and Lunch are mentioned more often on Sunday than on Monday.

There is essentially no discussion of meals during the period from 2 a.m. to 6 a.m. These plots also exhibit a small forward shift in time following Spring Forward, suggesting that each meal was tweeted about, and probably eaten, later in the day on Sunday. The effect is greater on the East Coast, and disappears on both coasts by Monday.

Broadening from messages mentioning specific meals to all messages, daily activity plots of  $B_{BSF}$  and  $B_{SF}$  reveal a regular diurnal pattern of behavior that is consistently shifted forward in time the evening following Spring Forward events. Figure 2 shows this shift for the year 2013, but the results were similar for other years.

Panel (a) suggests overall activity across the U.S. peaks around 9 p.m. on Sundays before Spring Forward (red circles), and experiences a minimum around 5am. The peak shifts approximately 45 min later on the Sunday of Spring Forward (blue squares) before synchronizing again by early morning Monday. In panel (b) California is used as an illustrative example of these patterns existing at the state level, and the smooth behavioral pattern constructed using Gaussian Process Regression. The pattern is similar to that observed for the entire country, with the exception of a slightly reduced amplitude. Twinflexion points are illustrated by black squares in panels (b) and (c).

Figure 2 demonstrates evidence that there is a shift in the peak time spent interacting with Twitter on Sunday evening following Spring Forward, relative to prior Sundays. Given the absence of a corresponding delay in interaction Monday morning, we infer a decrease in sleep opportunity experienced on Sunday night.



To explore the spatial distribution of the behavioral changes induced by Spring Forward, in Fig. 3 we map the time of peak Twitter activity on Sunday night for each state before (top) and the week of (bottom) Spring Forward, averaged across the years 2011–2014. On the Sundays leading up to Spring Forward (top), peak twitter activity occurs near either 10 p.m. for states on the East Coast, or 9:15 p.m., for most of the other states. The week of Spring Forward, nearly all states exhibit peak activity later in the night.

Looking at Texas as an individual example, before Spring Forward we see peak activity around 9:15 p.m. local time, and the week of Spring Forward it occurs at 10:15 p.m. local time. While Texas is one of the latest peaks observed on the evening following Spring Forward, several other states are up late as well including Oklahoma, Georgia, and Mississippi each peaking around 10:15 p.m.

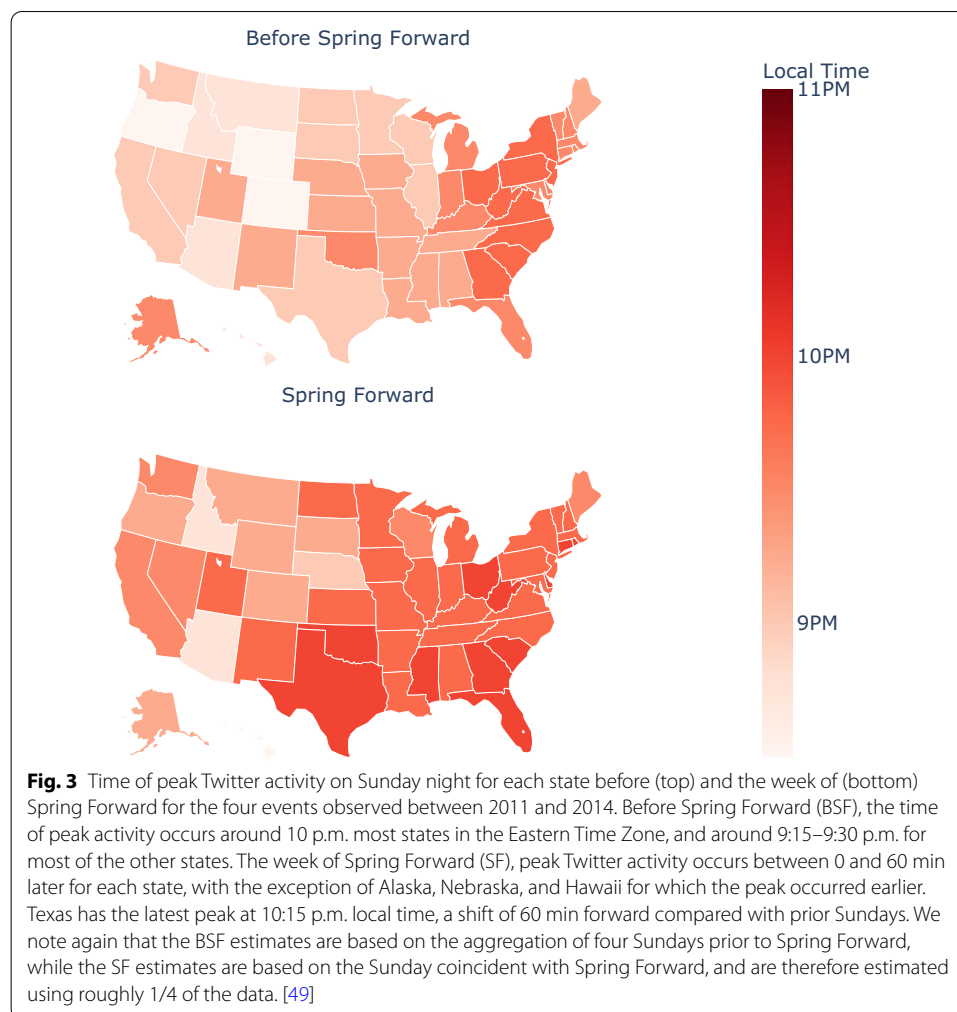
In the Additional file 1, we show maps estimating the time of peak activity for each of the individual 9 weeks centered on Spring Forward (see Additional file 1: Fig. S1). There is some week-to-week variation, most notably in the second week prior to Spring Forward, which was the night of the Academy Awards for three of the 4 years. By 4 weeks after Spring Forward, the peak activity map has relaxed to roughly the same pattern as BSF.

The magnitude of the forward shift in behavior illustrated in Fig. 3 is considered a proxy for the loss of sleep opportunity on the Sunday night following Spring Forward.

We used two distinct methods to estimate this magnitude, namely the peak shift and the twinflexion shift. A comparison of the spatial estimates made using each method are shown in Fig. 4.

Panel (a) illustrates the average shift in peak activity observed for 2011–2014 by computing the difference between the pair of maps in Fig. 3 (bottom minus top). There is



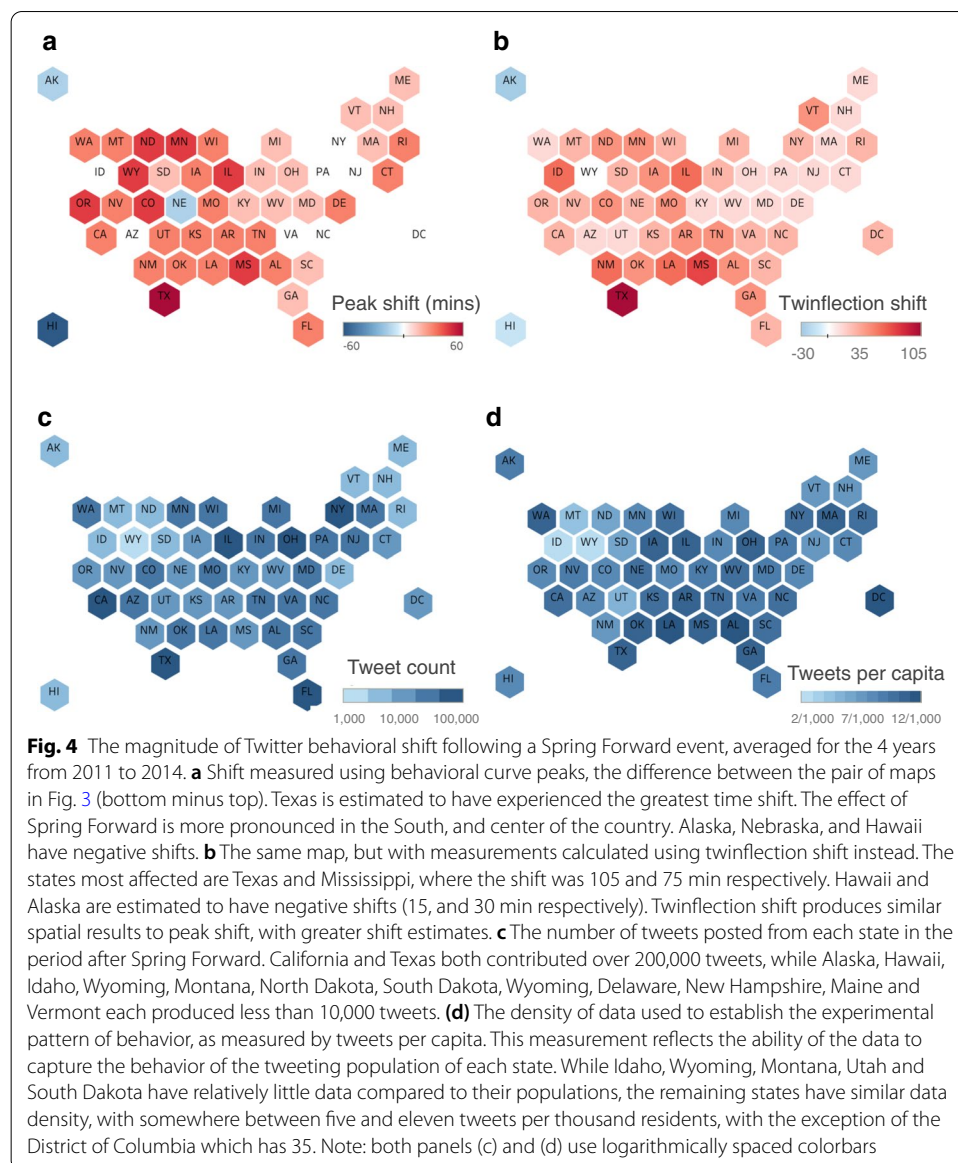


clear spatial variation in the shift in time on the night of Spring Forward, while most states exhibit a positive forward shift some exhibit none, and Alaska, Hawaii, and Nebraska show a negative shift. The peak in Twitter behavior for the east and west coasts occurred 0–30 min later Sunday night, while it occurred 30–60 min later for the central U.S. (Fig. 4a).

Figure 4b estimates the change using twinflexion, namely the change in concavity of the behavior activity curve from down to up. Every state except Hawaii, Alaska, and Wyoming exhibits a shift forward in time, and with similar spatial regularity. When measured with twinflexion shift, Texas and Mississippi are seen to have the greatest temporal shift following Spring Forward. Texans were tweeting 105 min later than usual following a Spring Forward event. Most of the east and west coast states were measured as tweeting 15 to 30 min later (Fig. 4b). Both measures agreed on a positive shift for the country as a whole. However, the two measures yielded different results for the magnitude of these shifts, with twinflexion shift generally estimating a more positive shift.

Figure 4c, d illustrate the amount of data contributing to calculations for the behavioral curves, and the density of this data with respect to each state's population. Idaho, Alaska, Hawaii, Montana, Wyoming, North Dakota, South Dakota, and Vermont were

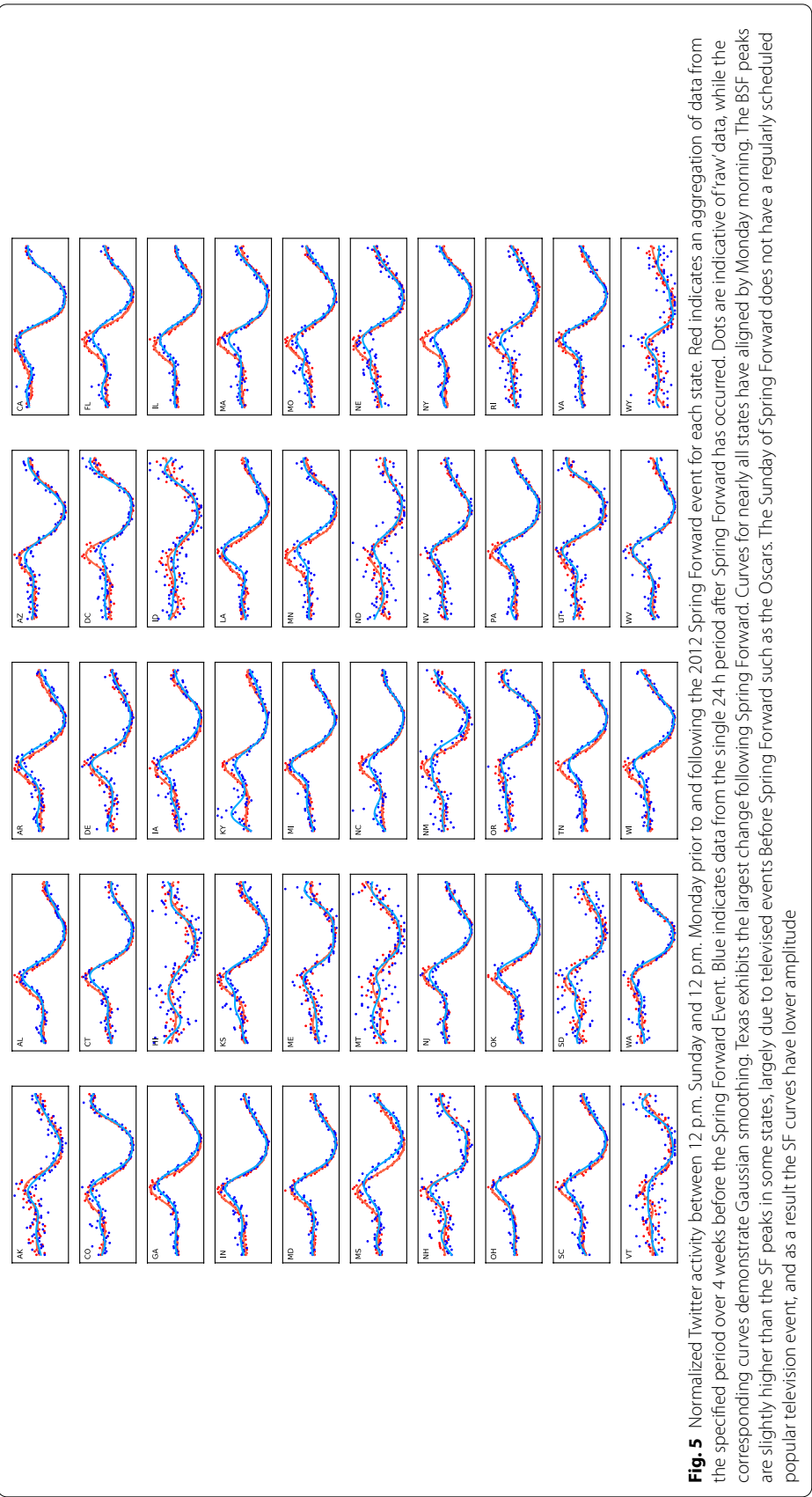


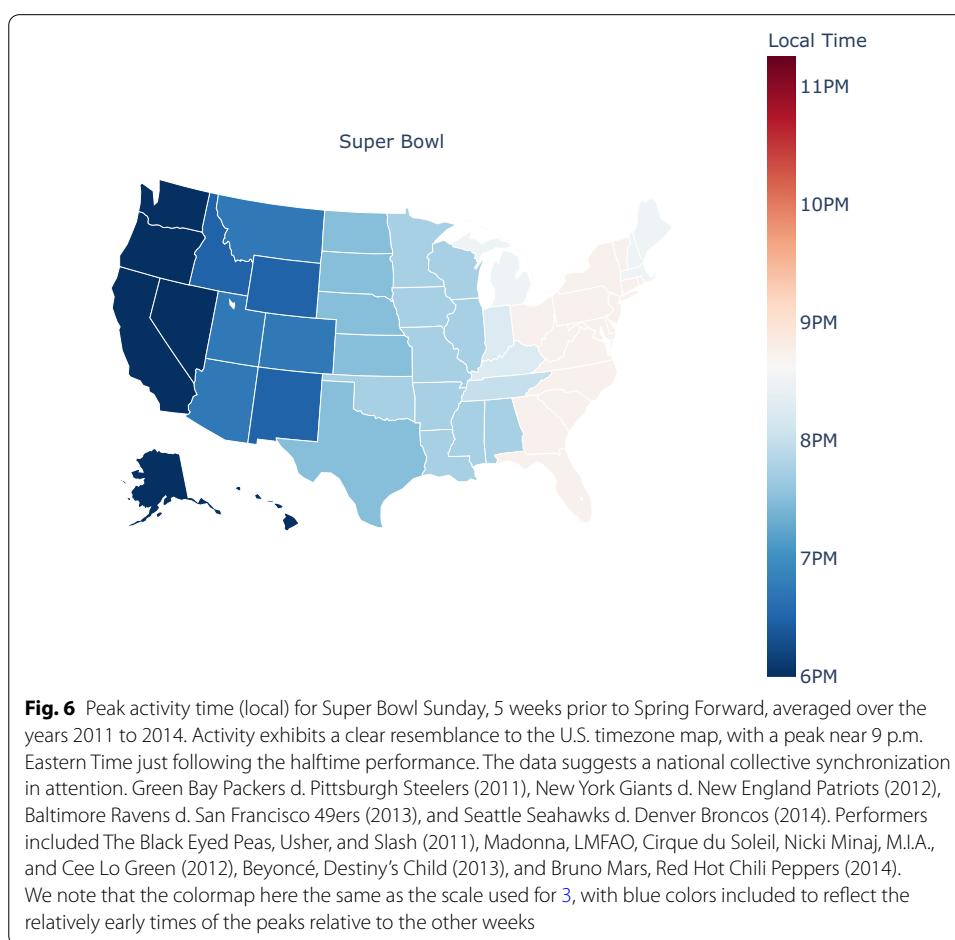


the states offering the smallest amount of data, and subsequently have the highest potential for a poor behavioral curve model fit. Wyoming was unique in that in 2013 for the 24 h observation window on the week of Spring Forward there were no tweets meeting inclusion requirements, making conclusions about this state particularly tenuous.

Though the amount of data available for California and Texas is much greater than the other states, when considering their large population size we find their twitter activity per capita to be similar to most other states. Based on our estimate of tweets per capita, we expect behavioral curves for most states to be more or less equally representative of their tweeting populations.

Looking at the diurnal cycle of Twitter activity for each individual state, we see remarkable consistency. Figure 5 shows the 24 h period spanning noon Sunday to noon Monday local time for the year 2012. Plots for the other 3 years exhibit similar behavior.





Before Spring Forward (red), most states show a peak between 9:15 and 10:00 p.m., local time. The week of Spring Forward (blue), nearly all states have a peak after 9:30 p.m. While states differ slightly in the time of peak, and magnitude of shift in the peak, most exhibit a clear positive shift (see Additional file 1: Fig. S3). By Monday morning, nearly all curves have re-aligned. We also consistently observe higher peaks for the BSF curves which we believe to be driven by televised events such as the Oscars. The Sunday of Spring Forward does not have a regularly scheduled popular television event, and as a result the SF curves have lower amplitude.

Both the peak and twinflexion demonstrate that it is possible to observe a measurable decrease in the amount of sleep opportunity people in the United States receive on average due to Spring Forward. They also both demonstrate uneven geographic distribution of the effect of Spring Forward, and therefore the ability to determine geographic disparity in sleep loss.

We also discovered that the Super Bowl occurred exactly 5 weeks prior to Spring Forward in each of the years studied. This annual event watched by over 100 million individuals in the U.S. caused peak Twitter activity to synchronize at roughly the same time nationally, around 9 p.m. Eastern, during the second half of the football game. The map in Fig. 6 shows the time of peak activity for each state on Super Bowl Sunday, averaged

over the years 2011 to 2014. The colormap is the same as the scale used for 3, with the additional cooler range brought in capture the time of peak relative to the usual times.

The map bears a remarkable resemblance to the timezone map, demonstrating a synchronization of collective attention across the country. Data from Super Bowl Sunday was not included in the Before Spring Forward data, as it does not accurately reflect the spatial distribution of typical posting behavior on a Sunday evening.

## Discussion

Technically speaking, Spring Forward occurs very early Sunday morning, and the instantaneous clock adjustment from 2 a.m. to 3 a.m. is witnessed by very few waking individuals. In addition, we speculate that the majority of individuals do not set an alarm clock for Sunday morning. As a result, we expect that the hour lost to Spring Forward will be felt by our bodies most meaningfully on Monday morning. Indeed, we are likely to experience the Monday morning alarm as occurring an hour early, as Spring Forward shortens the time typically reserved for sleep opportunity Sunday night by 1 h.

Considering the correlation between screen time and lack of sleep, the Sunday evening shift, and the corresponding Monday morning re-synchronization, we observe evidence that sleep opportunity is lost in some states on the evening of Spring Forward. By estimating the magnitude and spatial distribution of the shift in Twitter behavioral curves, we have approximated a lower bound on sleep loss at the state level.

Our pair of measurement methodologies have a Pearson correlation coefficient of 0.575, and a Spearman correlation coefficient of 0.467 (see Additional file 1: Fig S3). While they produced slightly different estimates of the magnitude of temporal shift in behavior, the resulting geographic profiles of sleep loss were similar. Both suggest that states along the coast are least affected by Spring Forward, while Texas and the states surrounding it to the North and East are the most affected.

Peak shift suggests the temporal shift in behavior due to Spring Forward generally less than the actual clock shift (1 h). California, the state for which we have the most data and therefore the most representative behavior profile after smoothing, was found to have a peak shift of 30 min.

Considering the clock adjustment of exactly 1 h, both measurements are plausibly directly representative of sleep lost, however the differing magnitudes of the measurements indicate that future work should clarify the relationship between these measurements and actual shifts. Twinflexion measured similar shifts for most states, but for a few estimated larger effects. While California was measured as having the same 30 min shift, Texas, the state for which we have the second most data, was estimated by twinflexion to be delayed by an additional 45 min.

Twinflexion measured a small forward shift for the state of Arizona, which does not observe DST. This could indicate that the twinflexion method overestimates the behavioral shift. It is also possible that a shift in behavior could occur for residents of Arizona, as a result of their connections to those in neighboring states, and in their former timezone. In example, some residents likely work in bordering states, and are forced to observe DST, and some will likely engage in more online activity and discussion when their peers are present- those peers being initially established by a shared

time of activity. This we believe to be an important distinction between Arizona and Hawaii, which also does not observe DST.

Hawaii is measured to have gained sleep opportunity by both accounts. Lacking the observation of DST, neighboring states, and other states in the same timezone, it is plausible that behavior in Hawaii would be unlike any other state, and be more independent of behaviors in other states. However, Hawaii's results should be considered tentative at best, given the sparsity of data available. This sparsity of data and relative independence from other states is shared with Alaska, the other state with a measured sleep opportunity gain by both measures. Caution should likewise be extended to measurements ascribed to South Dakota, North Dakota, Wyoming, Idaho, Montana, Vermont, New Hampshire, Rhode Island, Delaware, and Maine. These states have smaller populations, less population density, and lower volume of tweets. As a result, the behavioral curves associated with these states are less reliable.

Discrepancies in available data were determined to be largely accounted for by differences in population. Thus, we expect results for each state (exclusive of those mentioned earlier) to be comparably reliable in their representation of sleep loss for the state as a whole.

Incremental future work in this area could investigate state specific sleep loss related to Spring Forward events, which would allow further clarification of the relationship between the magnitude of behavioral shifts on Twitter and population sleep loss. Other directions might include looking at other sleep opportunity interruption events such as the end of Daylight Savings in November, where we are ostensibly given an additional hour of sleep opportunity. Our findings suggest that the sleep behavior associated with other annual events including New Year's Eve and Thanksgiving ought to be visible through tweets. This and other works would also benefit from exploration of the relationship between measurements of sleep opportunity as given by social media activity and actual sleep duration. More ambitiously, proxy data such as this could be verified by matching wearable measurements of sleep (e.g. Fitbit) with social media accounts.

Our study suffers from several limitations associated with our data source, we describe a few such examples here. The geographic location users provide in their Twitter bio is static and unlikely to be updated when traveling. As a result, user locations (time zone, state) inferred from this field will not always reflect their precise location. The GPS tagged messages included in our analysis will not suffer from this same uncertainty. Furthermore, the tweeting population of each state is likely to have complicated biases with respect to their representation of the general population [50].

Our dataset likely contains automated activity. Indeed, an entire ecology of algorithmic tweets evolved during the period in which we collected data for this study. However, we expect the majority of this activity to be scheduled using software that updates local time automatically in response to Daylight Savings. As such, this 'bot' type activity should largely serve to reduce our estimate of the time shift exhibited by humans.

As we showed for the Super Bowl, live televised events (e.g. sports, awards shows) have the potential to be a forcing mechanism to synchronize our collective attention throughout the week, and especially on Sunday evenings. Indeed, many individuals take to Twitter as a second screen during such events to interact with other viewers. In

addition, streaming services such as Netflix and HBO often release new episodes of popular shows on Sunday night to align with peak consumption opportunity. These cultural attractions exert a temporal organizing influence on our leisure behavior, and the Spring Forward disturbance translates this synchronization forward in time.

It is worth noting that early March is a rather dull time of year for popular professional sports in the United States. While the National Basketball Association and National Hockey League are finishing up their regular seasons, the National Football League is in its off-season and Major League Baseball beginning pre-season exercises. Arguably the most engaging live-televised sporting contests taking place in early March are the NCAA College Basketball Conference Championship games, with March Madness happening weeks after Spring Forward.

In 2014, the Academy Awards were hosted by Ellen DeGeneres on Sunday March 2. Her famous selfie tweet containing many famous actors was posted that evening, a message which held the record for most retweeted status update for several years [51]. The event happened the week before Spring Forward, and led to anomalous behavior compared with all other Sundays we looked at.

Since Spring Forward only occurs once per year, the specific language of the tweets is highly dependent on events occurring on that specific day. The variability in daily events and susceptibility of affect to these daily events makes study of the actual language in the tweets unreliable.

Finally, Twitter (and other social media companies) have access to much higher fidelity information regarding user activity than we have analyzed here. We are not able to analyze consumption activity on the site, e.g. when individual messages are interacted with via views, likes, or clicks. These forms of interaction with the Twitter ecosystem are likely to occur chronologically following the final posting of a message in the evening, and prior to the initial posting of a message in the morning. As a result, we expect our estimate of the sleep opportunity lost due to Spring Forward to be a lower bound.

## Conclusion

Privacy preserving passive measurement of daily behavior has tremendous potential to transform population-scale human activity into public health insight. The present study leverages a natural experiment in sleep loss to identify behavioral adaptation from Twitter data. It demonstrates a proof-of-concept along the path to a far more ambitious goal: construction of an ‘Insomniometer’ capable of real-time estimation of large-scale sleep duration and quality. Which cities in the U.S. slept well last night? Which states are increasingly suffering from insomnia? Answers to questions like these are not available today, but could lead to better public health surveillance in the near future. For example, communities exhibiting disrupted sleep in a collective pattern may be in the early stages of the outbreak of the flu or some other virus. Current methodologies for answering these questions are not scalable, but social media, mobile devices, and wearable fitness trackers offer a new opportunity for improved monitoring of public health.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40537-021-00503-0>.

**Additional file 1.** Additional figures S1–S3 and tables S1–S3.

### Acknowledgements

KL, TA, PSD, and CMD thank MassMutual for contributing funding in support of this research. The authors thank Adam Fox, Marc Maier, Jane Adams, David Dewhurst, Lewis Mitchell, and Henry Mitchell for helpful conversations.

### Authors' contributions

Research question was formulated by KL, CMD, PSD, and JL. Data manipulations were performed by KL and MA. Figures were generated by KL. Data analysis was performed by KL, and CMD. Writing was done by KL, TM, and CMD. Revisions were made by KL, TA, TM, CMD, PSD, and JL.

### Funding

KL, CD, and PSD were supported by a gift from MassMutual. JL is employed by MassMutual. MassMutual did not participate in the design of the study, or the collection, analysis, or interpretation of data.

### Data availability

Message counts used to construct behavioral curves are available at the paper's associated GitLab repository: (<https://gitlab.com/compsstorylab/insomniometer>). We used Twitter's gardenhose stream (10% of all tweets); we do not share messages for privacy reasons, and our data products and results are population-scale artifacts. Using samples of tweets from the same time periods will generate similar results.

## Declarations

### Ethics approval and consent to participate

Research using aggregate Twitter data was approved for IRB exemption by UVM under protocol 15-571: Building an Instrument that Measures Health, Well-Being, and Disease in Real-Time.

### Consent to participate

Not applicable.

### Competing interests

The authors are unaware of any competing interests at this time.

### Author details

<sup>1</sup>Computational Story Lab, Vermont Complex Systems Center, MassMutual Center of Excellence for Complex Systems & Data Science, University of Vermont, Burlington, USA. <sup>2</sup>Lehigh University, Bethlehem, USA. <sup>3</sup>MassMutual Data Science, Amherst, USA.

Received: 8 April 2021 Accepted: 20 August 2021

Published online: 09 September 2021

## References

1. Panel CC, Watson NF, Badr MS, Belenky G, Bliwise DL, Buxton OM, et al. Joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society on the recommended amount of sleep for a healthy adult: Methodology and discussion. *Sleep*. 2015;38(8):1161–83.
2. Short Sleep Duration Among US Adults; 2017.
3. Ford ES, Cunningham TJ, Croft JB. Trends in self-reported sleep duration among US adults from 1985 to 2012. *Sleep*. 2015;38(5):829–32.
4. Althoff T, Horvitz E, White RW, Zeitzer J. Harnessing the web for population-scale physiological sensing: a case study of sleep and performance. In: WWW '17 Proceedings of the 26th International Conference on World Wide Web. 2017; p. 113–22.
5. Curcio G, Ferrara M, De Gennaro L. Sleep loss, learning capacity and academic performance. *Sleep Med Rev*. 2006;10(5):323–37.
6. Engle-Friedman M. The effects of sleep loss on capacity and effort. *Sleep Sci*. 2014;7(4):213–24.
7. Rosekind MR, Gregory KB, Mallis MM, Brandt SL, Seal B, Lerner D. The cost of poor sleep: workplace productivity loss and associated costs. *J Occup Environ Med*. 2010;52(1):91–8.
8. Dean B, Aguilar D, Shapiro C, Orr WC, Isserman JA, Calimlim B, et al. Impaired health status, daily functioning, and work productivity in adults with excessive sleepiness. *J Occup Environ Med*. 2010;52(2):144–9.
9. Owens J, Dingus T, Guo F, Fang Y, Perez M, McClafferty J, et al. Prevalence of drowsy-driving crashes: Estimates from a large-scale naturalistic driving study. AAA Foundation for Traffic Safety. 2018.



10. Anderson C, Ftouni S, Ronda JM, Rajaratnam SM, Czeisler CA, Lockley SW. Self-reported drowsiness and safety outcomes while driving after an extended duration work shift in trainee physicians. *Sleep*. 2018;41(2):zsx195.
11. Medic G, Wille M, Hemels ME. Short-and long-term health consequences of sleep disruption. *Nat Sci Sleep*. 2017;9:151.
12. Nagai M, Hoshida S, Kario K. Sleep duration as a risk factor for cardiovascular disease—a review of the recent literature. *Curr Cardiol Rev*. 2010;6(1):54–61.
13. Cheung V, Yuen V, Wong G, Choi S. The effect of sleep deprivation and disruption on DNA damage and health of doctors. *Anaesthesia*. 2019;74(4):434–40.
14. Fox M. Shift work may cause cancer, world agency says. Reuters. 2007.
15. Patel NP, Grandner MA, Xie D, Branas CC, Gooneratne N. “Sleep disparity” in the population: poor sleep quality is strongly associated with poverty and ethnicity. *BMC Public Health*. 2010;10(1):475.
16. Chattu VK, Chattu SK, Spence DW, Manzar MD, Burman D, Pandi-Perumal SR. Do disparities in sleep duration among racial and ethnic minorities contribute to differences in disease prevalence? *J Racial Ethnic Health Disparit*. 2019;6(6):1053–61.
17. Ruten ME, DeCoster J, Jacobs L, Lichstein KL. Normal sleep in African-Americans and Caucasian-Americans: a meta-analysis. *Sleep Med*. 2011;12(3):209–14.
18. Curtis DS, Fuller-Rowell TE, El-Sheikh M, Carnethon MR, Ryff CD. Habitual sleep as a contributor to racial differences in cardiometabolic risk. *Proc Nat Acad Sci*. 2017;114(33):8889–94.
19. Hafner M, Stepanek M, Taylor J, Troxel WM, Van Stolk C. Why sleep matters—the economic costs of insufficient sleep: a cross-country comparative analysis. *Rand Health Q*. 2017;6(4):11.
20. Bianchi MT. Sleep devices: wearables and nearables, informational and interventional, consumer and clinical. *Metabolism*. 2018;84:99–108.
21. Douglas NJ, Thomas S, Jan MA. Clinical value of polysomnography. *Lancet*. 1992;339(8789):347–50.
22. Harvey AG, Tang NK. (Mis) perception of sleep in insomnia: a puzzle and a resolution. *Psychol Bull*. 2012;138(1):77.
23. Lauderdale DS, Knutson KL, Yan LL, Liu K, Rathouz PJ. Self-reported and measured sleep duration: how similar are they? *Epidemiology*. 2008;19:838–45.
24. Marino M, Li Y, Rueschman MN, Winkelmann JW, Ellenbogen J, Solet JM, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep*. 2013;36(11):1747–55.
25. Jonasdottir SS, Minor K, Lehmann S. Gender differences in nighttime sleep patterns and variability across the adult lifespan: a global-scale wearables study. *Sleep*. 2021. <https://doi.org/10.1093/sleep/zsaa169>.
26. Roomkham S, Lovell D, Cheung J, Perrin D. Promises and challenges in the use of consumer-grade devices for sleep monitoring. *IEEE Rev Biomed Eng*. 2018;11:53–67.
27. Leypunskiy E, Kiciman E, Shah M, Walch OJ, Rzhetsky A, Dinner AR, et al. Geographically resolved rhythms in Twitter use reveal social pressures on daily activity patterns. *Curr Biol*. 2018;28(23):3763–75.
28. Christensen MA, Bettencourt L, Kaye L, Moturu ST, Nguyen KT, Olgin JE, et al. Direct measurements of smartphone screen-time: relationships with demographics and sleep. *PLoS ONE*. 2016;11(11):e0165331.
29. Rios M, Lin J. Visualizing the pulse of world cities on Twitter. In: Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media; p. 4.
30. Roenneberg T. Twitter as a means to study temporal behaviour. *Curr Biol*. 2017;27(17):R830–2. <https://doi.org/10.1016/j.cub.2017.08.005>.
31. Scheffler T, Kyba CCM. Measuring Social Jetlag in Twitter Data. In: Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016); p. 4.
32. Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS ONE*. 2011;6(12):e26752.
33. Alajajian SE, Williams JR, Reagan AJ, Alajajian SC, Frank MR, Mitchell L, et al. The Lexicocalorimeter: Gauging public health through caloric input and output on social media. *PLoS ONE*. 2017;12(2):e0168893.
34. Dzugog F, Lightman S, Cristianini N. Circadian mood variations in Twitter content. *Brain Neurosci Adv*. 2017;1:2398212817744501. <https://doi.org/10.1177/2398212817744501>.
35. Dzugog F, Lightman S, Cristianini N. Diurnal variations of psychometric indicators in Twitter content. *PLoS ONE*. 2018;13(6):e0197002. <https://doi.org/10.1371/journal.pone.0197002>.
36. McIver DJ, Hawkins JB, Chunara R, Chatterjee AK, Bhandari A, Fitzgerald TP, et al. Characterizing sleep issues using Twitter. *J Med Internet Res*. 2015;17(6):e140.
37. Martín-Olalla JM. The long term impact of daylight saving time regulations in daily life at several circles of latitude. *Sci Rep*. 2019;9(1):1–13.
38. Martín-Olalla JM. Scandinavian bed and rise times in the Age of Enlightenment and in the 21st century show similarity, helped by Daylight Saving Time. *J Sleep Res*. 2019. <https://doi.org/10.1111/jsr.12916>.
39. Kantermann T, Juda M, Merrow M, Roenneberg T. The Human Circadian Clock's seasonal adjustment is disrupted by daylight saving time. *Curr Biol*. 2007;17(22):1996–2000. <https://doi.org/10.1016/j.cub.2007.10.025>.
40. Sandhu A, Seth M, Gurm HS. Daylight Savings Time and myocardial infarction. *Open Heart*. 2014;1(1):e000019.
41. Sipilä JO, Ruuskanen JO, Rautava P, Kytö V. Changes in ischemic stroke occurrence following Daylight Saving Time transitions. *Sleep Med*. 2016;27:20–4.
42. Varughese J, Allen RP. Fatal accidents following changes in Daylight Savings Time: The American experience. *Sleep Med*. 2001;2(1):31–6.
43. Martín-Olalla JM. Traffic accident increase attributed to Daylight Saving Time doubled after Energy Policy Act. *Curr Biol*. 2020;30(7):R298–300. <https://doi.org/10.1016/j.cub.2020.03.007>.
44. Kamstra MJ, Kramer LA, Levi MD. Losing sleep at the market: The Daylight Saving anomaly. *Am Econ Rev*. 2000;90(4):1005–11.

45. Gray TJ, Reagan AJ, Dodds PS, Danforth CM. English verb regularization in books and tweets. *PLoS ONE*. 2018;13(12):e0209651.
46. Twitter. Developer application program interface (API); 2020. <https://developer.twitter.com/en/docs/tweets/sample-realtime/overview/decahose>.
47. Rasmussen CE. Gaussian processes in machine learning. In: summer school on machine learning. Springer; 2003. p. 63–71.
48. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12(Oct):2825–30.
49. Inc PT. Collaborative data science; 2015. <https://plot.ly>.
50. Wojcik S, Hughes A. Sizing up Twitter users; 2019.
51. DeGeneres E; March 2, 2014. <https://twitter.com/theellenshow/status/44032224407314432>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---