Real, misreported, and backfilled adherence with paper sleep diaries

Megan Clegg-Kraynok a, Lauren Barnovsky b, Eric S. Zhou c, d, e, *

a Department of Psychology, Ohio Northern University, 525 South Main St, Ada, OH, 45810, USA
b Department of Criminology and Criminal Justice, University of Maryland, 7251 Preinkert Dr, College Park, MD, 20742, USA
c Division of Sleep Medicine, Harvard Medical School, 221 Longwood Avenue, Boston, MA, 02115, USA
d Perini Family Survivors’ Center, Dana-Farber Cancer Institute, 450 Brookline Ave, Boston, MA, 02215, USA
e Department of Neurology, Boston Children’s Hospital, 300 Longwood Ave, Boston, MA, 02115, USA

Article info
Article history:
Received 5 February 2023
Received in revised form 12 April 2023
Accepted 13 April 2023
Available online 14 April 2023

Keywords:
Sleep diary
Adherence
Insomnia disorder

Abstract
Objective: Paper-based sleep diaries play an important role in the diagnosis and treatment of insomnia disorder. Accurate self-report data help to guide therapy and track progress, both in the clinic and during research trials. Previous research with paper diaries suggests that timely adherence with self-report diaries may be an issue, which can result in biased event recall.

Patients/methods: University students (N = 31) were asked to track their bedtime and wake time within 30 min of these events on paper-based sleep diaries. Specially designed binders covertly timestamped when participants actually wrote on their sleep diary. We assessed adherence by comparing time-stamped diary usage with what participants documented in their sleep diary.

Results: Participants self-reported they were adherent with sleep diary instructions 97.9% of the time. However, timestamped data revealed that only 37.1% of diary entries were completed within the instructed timeframe. More than half of participants backfilled diary data, and three participants (9.7%) provided data that completely did not match their actual time of completion.

Conclusions: When naïve to the objective tracking of their sleep diary usage, participants greatly over-reported the extent of their adherence. Non-adherence with sleep diary protocols poses a challenge for researchers utilizing this tool as a study outcome in clinical trials and for clinicians attempting to implement behavioral therapies for insomnia.

© 2023 Elsevier B.V. All rights reserved.

1. Introduction

An estimated 50 to 70 million people in the U.S. meet diagnostic criteria for at least one sleep disorder [1,2]. The most common is insomnia disorder, which is characterized by difficulty with falling and/or staying asleep, resulting in daytime sequelae. Approximately 10% of the adult population suffers from insomnia disorder [3]. When it remains untreated, insomnia is associated with a number of adverse outcomes including the development of depression [4], cardiovascular disease [5], and cancer [6].

Consistent and compelling data has demonstrated the effectiveness of behavioral and psychological treatments for insomnia disorder in adult populations, culminating in recommendations for cognitive-behavioral therapy for insomnia (CBT-I) as frontline treatment being made by the American Academy of Sleep Medicine [7] and American College of Physicians [8]. CBT-I is a multi-component protocol that includes sleep restriction, stimulus control, cognitive therapy, sleep hygiene, and relaxation therapy. As insomnia is a subjective disorder, diagnosed not by polysomnography or actigraphy but by patient report, the daily tracking of sleep via self-report sleep diaries is an important part of CBT-I. Sleep diaries not only provide patients and clinicians with information about relevant sleep metrics at baseline (e.g., sleep onset latency, wake after sleep onset, total time in bed), but is used to guide the implementation of sleep restriction, and then continuously throughout therapy to track the impact of treatment. Prior literature has demonstrated that patients who are more adherent with CBT-I homework (e.g., sleep diaries) benefit more from therapy [9]. Sleep diaries are so important in the treatment of insomnia disorder that there is widespread agreement that they should be routinely included in all insomnia-related research [10].

The phenomenon of backfilling self-report data or offloading...
medications just prior to an appointment is a known problem in medical trials [11]. There are a number of reasons why an individual may not be adherent with self-report diary directions, including managing competing demands for their time, simple forgetfulness with the introduction of a new behavioral task that is not part of a daily routine, and miscommunication with providers about its relevance to treatment goals. Stone and colleagues published a landmark study in which they objectively assessed adherence with paper diaries assessing pain by covertly capturing adherence as reported by participants and comparing this data to an electronic record of the event [12]. Results showed that participants claimed to be adherent with study instructions 90% of the time, but were shown to be adherent on a mere 11% of diary entries, calling “into question the use of paper diaries.” In a well-designed study examining the issue of sleep diary adherence, a population of non-clinical young adults were asked to follow a specific sleep schedule. Those who were not told that the actigraphy they were wearing would be used to track both their sleep as well as their adherence with the schedule recommendations reported following the researcher’s directions on self-report sleep diaries, with the objective actigraphy data telling a different tale [13].

Metrics calculated from a patient’s sleep diaries directly inform treatment recommendations. Despite the importance of accurate data, we are not aware of existing research that has studied objective adherence with paper sleep diaries with such careful attention paid to concealing the fact that participants were being monitored. Research has demonstrated that human memory is fallible, with our recall for prior events influenced by cognitive biases [14]. As there are a number of sleep misperceptions common among patients with insomnia disorder [15], it is imperative that we understand the extent to which paper sleep diary data is reliably documented.

2. Methods

The single-center study was conducted at a private university in the U.S. Data was collected between September 2020 and April 2022. Recruitment into the study was paused between October 2020 to February 2021 due to COVID restrictions mandated by university administration. The study protocol was approved by the IRB at Ohio Northern University and all participants provided informed consent.

2.1. Participants

A convenience sample of young adults, at least 18 years of age, were recruited from lower-level university psychology courses and enrolled into the study. Participants were not required to present with an active sleep disorder. They were offered course credit for study participation, with alternative means of earning equal credit provided to avoid coercion to participate. Of the 53 participants enrolled, data from 31 were included in analyses, with the remainder excluded due to equipment malfunction. The issues experienced were insufficient battery life that failed that last a full week (n = 12), failure to record any data on the SD card (n = 5), the device resetting the internal clock to a default date/time that would not allow for accurate data interpretation (n = 4), and failure to return the equipment to investigators (n = 1). There were no significant differences comparing those included in data analyses versus those excluded on all study demographic variables seen in Table 1.

2.1.1. Objective sleep diary adherence tracking

Binders were specifically designed for this study by a medical device product development and manufacturing company in Tyngsboro, MA. First, small metal binders (3” x 7”) were purchased from an office supply store. On the inside of each binder, a pressure sensor capable of detecting when someone wrote on a piece of paper was installed, with each event timestamped and stored on a memory card that could be downloaded via USB. A 3D-printed housing fully concealed these electronic components. Preprinted paper sleep diaries were then placed on top of the pressure sensor, held in place by a clip at the top of the binder. To decrease the likelihood of a participant accidently triggering the pressure sensor, the binder included a spring mechanism that automatically closed each time that it is opened. Five binders were produced, with each unit field-tested by study investigators for a period of at least one week. Additionally, each unit was tested prior to use by a study participant (See Fig. 1).

2.1.2. Procedure

Prior to study consent, participants were offered the opportunity to take part in a study assessing sleep in college students and would be asked to track sleep variables using a sleep diary for two weeks. Once consented, participants were provided with paper-based sleep diaries stored in a study binder. They were instructed to track the time they went to bed at night, the time they woke up in the morning, the time of any daytime naps, and the number of night wakeings within 30 min of their occurrence. If a participant recorded these times outside the prescribed 30-min window, they were asked to indicate this with an asterisk by that recording. They were also instructed to leave their paper sleep diaries in the binder provided to avoid losing them. Participants were naïve to the fact that the binders containing the sleep diaries were timestamping when they wrote on their sleep diaries. During consent, participants were instructed that the purpose of the study was to monitor sleep among emerging adults. Upon study completion, all participants were debriefed and informed that the purpose of the study was to examine participant adherence to research protocols. They were told that the binders included technology that time-stamped any time they opened or closed the binders and any time they wrote on the sleep diaries, and that if they had any issues with the use of their data, they could withdraw from the study at that point while still receiving full course credit. No study participant provided negative feedback about the study procedures or requested study withdrawal.

2.1.3. Data cleaning

The batteries used in the study binders needed to be small in order to fit without attracting a participant’s attention. Combined with battery degradation over time and repeated recharging cycles, insufficient battery life resulted in battery drainage before the end

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Sample demographics (N = 31).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD) or n (%)</td>
</tr>
<tr>
<td>Age</td>
<td>19.2 (1.2)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>21 (67.7%)</td>
</tr>
<tr>
<td>Male</td>
<td>9 (29.0%)</td>
</tr>
<tr>
<td>Non-binary</td>
<td>1 (3.2%)</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>2 (6.5%)</td>
</tr>
<tr>
<td>Latinx</td>
<td>1 (3.2%)</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>1 (3.2%)</td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>27 (87.1%)</td>
</tr>
<tr>
<td>Class Standing</td>
<td></td>
</tr>
<tr>
<td>Freshman</td>
<td>17 (54.8%)</td>
</tr>
<tr>
<td>Sophomore</td>
<td>5 (16.1%)</td>
</tr>
<tr>
<td>Junior</td>
<td>4 (12.9%)</td>
</tr>
<tr>
<td>Senior</td>
<td>5 (16.1%)</td>
</tr>
</tbody>
</table>
of the two-week data collection. This affected 18 of the 31 study participants. Consistent with clinical practice and many research trials, 7 days of data for each participant were included in our analyses. Additionally, the internal clocks of several binders inadvertently reset to a baseline date and time. For these datasets, the binder data was compared against sleep diary times to manually adjust to the correct time stamps. This was completed by computing the time elapsed between binder data points and between the reported sleep diary times. Three or more matching durations were found to determine an approximation of the real time and the data was adjusted accordingly. Additionally, the adjusted time stamp was compared to the date and time the participant received the binder to ensure that there was corresponding data. Due to the nature of the research question, there was concern that participants could have simply been consistently entering their data a certain duration following the reported time. However, comparisons between the results of real data and adjusted data did not show any significant differences in the means.

Finally, one additional sleep diary question was asked about caffeine consumption. Due to inconsistencies in participant completion of this item, we did not analyze this data.

2.1.4. Statistical analysis

Standards for real adherence, reported adherence, and misreported adherence were derived from prior research conducted on adherence with paper diaries [12]. **Real** adherence was defined by the number of times a participant wrote on their sleep diary, as tracked by the binder, within two fixed windows: ±30 min of when they wrote that they went to bed, woke up, or took a nap in the sleep diary (per study instructions; e.g., between 10PM and 11PM if a participant said they went to bed at 10:30PM) and ±60 min (to provide participants with a grace period). **Reported** adherence was calculated by counting the number of times a participant marked an asterisk next to their entry indicating that they had not entered the information within 30 min of it occurring, and then dividing this by the total number of diary entries completed. **Misreported** adherence was the difference between a participant’s reported adherence and their real adherence. **Backfilled** adherence, when a participant does not complete data as instructed, but backfills the data at a different time, was determined based on the number of days that they did not write in their sleep diary, suggesting that they retrospectively completed their diaries. Days in which both the wake and sleep times were marked by an asterisk were not counted as an occurrence of backfilled adherence. Finally, **minimal acceptable** adherence threshold was based on whether a participant was tracked as actually writing in their diary on at least 50% of their entries within a ±30-min window of the time that a sleep-related event was noted in their diary.

We computed the mean and 95% confidence intervals for all of the adherence outcomes. Testing for group differences in terms of real and minimal acceptable adherence based on demographic characteristics of gender and class standing used t-tests. Due to the small sample size and nature of primarily receiving participants from introductory courses, a binary variable of freshman versus sophomores and above was created for class standing. Stata version 17 was used for data analysis.

3. Results

Study participants were an average of 19.2 years of age (SD = 1.2), with more females than other genders (67.7%). They were predominantly non-Hispanic White (87.1%) and freshman (54.8%). This data is seen in Table 1.

3.1. Sleep diary adherence

**Real**: Of the 434 diary entries that were completed by the 31 participants, 37.1% were documented in the sleep diary within a ±30-min window of a corresponding sleep event (refer to Table 2). When increasing this time to a ±60-min window, real adherence reaches 45.4%. One participant completed all of their entries within the requested 30-min window. **Reported**: These figures differ from
the reported adherence of 97.9%. Misreported: Thus, for the 30- and 60-min windows, 60.8% and 52.5% of the total entries were misreported, respectively. Three study participants (9.7%) misreported all of their data. Backfilled: There were a total of 186 days where sleep diary data were reported. On 35 of these days (18.8%), study participants did not write on their sleep diary at all, but still documented at least one sleep event on that day. A total of 16 participants (51.6%) backfilled data on at least one day. Finally, 11 out of 31 participants (35.5%) demonstrated minimal acceptable adherence. This data is shared in Table 2.

4. Discussion

Sleep diaries are an important component in the management of insomnia disorder. They serve as the primary (and, sometimes only) data source for a clinician’s therapeutic recommendations and are an important outcome in many research trials. Sleep diary data must accurately reflect the patient’s perceptions of their insomnia symptoms for a treatment plan to be effective and to ascertain the efficacy of an intervention protocol. Because of the notable biases present in retrospective self-reports, which increase with more time between an event and its documentation, it is important to understand the extent to which paper-based sleep diaries are routinely completed within a reasonable timeframe. Our novel findings demonstrate that the majority of paper-based sleep diaries entries may be misreported, even though study participants claimed almost perfect adherence. Though we are entering an increasingly ‘digital age,’ the lessons we learned from this study on paper sleep diaries are ones that still may apply to electronic sleep diaries (e.g., app/web-based) as well.

Non-adherence with sleep diaries is not surprising. They represent a burden for patients, with the Consensus Sleep Diary asking 10 questions on a daily basis [16]. In our study, we asked participants for far less. This intentional design was meant to assess the bare minimum of sleep events, with the rationale that if someone were not able to complete these tasks as instructed, they would be unlikely to do more. Even with our minimum threshold for participant burden, and after relaxing our standards from a ±30-min window to a ±60-min window, we see that the 95% confidence interval for real adherence is between 34 and 57%. Thus, even in the most optimistic of scenarios, over 40% of sleep diary data may be a product of biased recall. This lack of adherence is a consistent dilemma across all study participants, with all but one individual misreporting at least one datapoint. It has been demonstrated that those who intentionally provide false responses affect outcomes in research trials much more than those who simply made a mistake in self-report data [17]. Researchers conducting clinical trials involving patients with insomnia should consider misreported sleep diary data as a very real possibility that can affect study results and be cautious about only using sleep diary completion as a metric of program adherence [18].

With advancing technological capabilities and decreasing costs, there is interest in the use of objective sleep trackers to passively collect sleep metrics as a way to possibly circumvent the issue of adherence with paper-based diaries altogether. However, there are some circumstances which necessitate the use of paper diaries (e.g., a patient population with more limited technology access or capabilities). While there is a need for much more research assessing the validity and reliability of consumer-facing technologies [19], there is increasing interest in the use of these devices in the management of insomnia [20,21]. In the work that has been conducted, data suggests that there are discrepancies between self-reported paper (and tablet-based) sleep diaries with outcomes assessed using actigraphs and consumer wearables [22–24]. Interestingly, it appears that the active task of keeping a sleep diary, and possibly combining this with an objective tracking device, could benefit patients. First, it was demonstrated that participants in a digital CBT-I intervention whose sleep was passively tracked by a wearable device, rather than the use of sleep diaries, reported worse sleep than those who did not [25]. Next, when participants were made aware that an objective tracker is monitoring their sleep, they were more likely to be honest on sleep diaries and to adhere with treatment recommendations [13]: Clinicians treating patients with insomnia are encouraged to: (1) consider supplementing paper sleep diaries with an objective tracker; (2) dedicating additional time during initial visits to ensure that patients understand how to complete the sleep diary and fully appreciate the importance of accurate data during their treatment course; and/or (3) electronic sleep diary data collection. Researchers who utilize digital sleep diaries that provide a data entry timestamp should consider assessing adherence as part of their study analyses, including assessing the extent to which adherence to sleep diaries may be linked to treatment outcomes.

4.1. Study limitations

Our novel research is not without limitations. First, we encountered technological issues with our sleep diary binders. These problems resulted in failed data collection, extra data cleaning, and analyses that examined data over a single week rather than two. Despite these challenges, the consistent patterns seen in our results suggests that the signals we identified are important ones to study further. Second, the modified sleep diary that was used in our study had not undergone a full validation study. Thus, adherence rates may have been impacted by participant comprehension difficulties and future research could consider the use of validated sleep diaries (e.g., the Consensus Sleep Diary). Third, our study population was comprised of adults who were not actively seeking treatment for insomnia disorder. It is conceivable that patients with insomnia may be more likely to consistently track their sleep because of the perceived benefits of treatment engagement. While conducting this study in the context of a clinical trial for insomnia disorder would improve generalizability, the use of deception in such a study could potentially harm the patient’s trust in providers offering insomnia interventions. Data from medical trials demonstrates non-adherence behaviors similar to our findings. For example, in one trial studying inhaler use to improve pulmonary function, only 15% of participants followed study directions when assessed objectively, compared to 73% who self-reported accurate use [11]. In addition to wearables, there are promising new efforts to collect sleep metrics via smart speakers as part of interventions, all of which could provide additional data to study adherence with sleep diaries in patients with insomnia disorder [26]. Fourth, we do not know the extent to which sleep diary data is misrepresented. As each timestamped event is only associated with when a participant wrote on the paper diary, we cannot
be certain of what they wrote. Further, it is possible that participants may have documented their sleep outcomes separately (e.g., on their phone) within the instructed window and transferred this information at a later time to their diaries. In addition, some sleep diary data may be less susceptible to recall bias than others (e.g., wake time as it could be linked to a morning alarm). Future work to determine the impact of missing an adherence window on the veracity of the data, as well as to compare paper-based and digital-based sleep diary adherence will be important.

4.2. Conclusion

Without an objective assessment of adherence, both clinicians and researchers should view paper sleep diary data with a wary eye. Clinicians must consider the possibility that the sleep metrics they are basing treatment decisions on could be impacted by the patient’s perceptions of their sleep health to a greater extent than they currently believe. Researchers may wish to exercise caution if choosing to only use paper-based sleep diaries as a primary study outcome measure, and consider alternative data collection methods (e.g., digital sleep diaries with data entry timestamps) in future trials. When it comes to intervention trials, it may be important to consider incentivizing accurate paper-based sleep diary data collection as part of the study protocol [27].

Funding details

Funding for this project was generously provided by a Department of Pediatric Hematology/Oncology Population Sciences research grant and Dr. Jennifer W. Mack.

CRediT authorship contribution statement

Megan Clegg-Kraynok: Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. Lauren Barnovsky: Formal analysis, Investigation, Data curation, Writing – review & editing. Eric S. Zhou: Conceptualization, Methodology, Validation, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Funding acquisition.

Declaration of competing interest

Eric S. Zhou has received grant support from Jazz Pharmaceuticals and Harmony Biosciences, and consulting fees from Samsung and MindUP for work unrelated to this manuscript.

Acknowledgments

The novel sleep diary technology would not have been possible without the ingenuity of a team at Northeast Biomedical (Timothy M. Looney, Philip Andrusin, and Angela Valente) who made magic happen with some sensors and a 3D printer. Finally, Dr. Arthur A. Stone was beyond kind with his time in discussing his prior research with patient adherence.

References


