

RESEARCH

Open Access



# An automatic estimation of the rest-interval for MotionWatch8© using uniaxial movement and lux data

Daniel Backhouse<sup>1,2</sup>, Ryan Stanley Falck<sup>1,2</sup> and Teresa Liu-Ambrose<sup>1,2\*</sup>

## Abstract

**Background:** Poor sleep is linked with chronic conditions common in older adults, including diabetes, heart disease, and dementia. Valid and reliable field methods to objectively measure sleep are thus greatly needed to examine how poor sleep impacts older adults. Wrist-worn actigraphy (WWA) is a common objective measure of sleep that uses motion and illuminance data to estimate sleep. The rest-interval marks the time interval between when an individual attempts to sleep and the time they get out of bed to start their day. Traditionally, the rest-interval is scored manually by trained technicians, however algorithms currently exist which automatically score WWA data, saving time and providing consistency from user-to-user. However, these algorithms ignore illuminance data and only considered motion in their estimation of the rest-interval. This study therefore examines a novel algorithm that uses illuminance data to supplement the approximation of the rest-interval from motion data.

**Methods:** We examined a total of 1086 days of data of 129 participants who wore the MotionWatch8© WWA for  $\geq 14$  nights of observation. Resultant sleep measures from three different parameter settings were compared to sleep measures derived following a standard scoring protocol and self-report times.

**Results:** The algorithm showed the strongest correlation to the standard protocol ( $r = 0.92$  for sleep duration). There were no significant differences in sleep duration, sleep efficiency and fragmentation index estimates compared to the standard scoring protocol.

**Conclusion:** These results suggest that an automated rest-interval scoring method using both light exposure and acceleration data provides comparable accuracy to the standard scoring method.

**Keywords:** Actigraphy, Sleep, Automatic, Rest-interval

## Introduction

Poor and insufficient sleep is associated with a number of health risks in older adults including diabetes, heart disease, dementia, and immune disorders (Gottlieb et al. 2006; Kasasbeh et al. 2006; Opp and Toth 2003; Shokri-Kojori et al. 2018). Reliable and valid field-methods for objectively measuring sleep quality are thus important for understanding these relationships and improving patient outcomes.

While polysomnography is the “gold-standard” in assessing sleep quality (Marino et al. 2013), it is resource intensive and requires subjects to be monitored overnight using various equipment. This makes it challenging to assess an individual's natural sleep-wake cycle using polysomnography. Self-report measures are a cheap, low-burden alternative to polysomnography which can capture people's natural sleep patterns over several days, weeks, or months (Matthews et al. 2018); however, they are open to reporting errors and individual biases (Lauderdale et al. 2008; Short et al. 2013).

Wrist-worn actigraphy (WWA) aims to reduce reporting errors and self-report bias by providing a non-invasive objective field-method for capturing sleep over

\* Correspondence: [teresa.ambrose@ubc.ca](mailto:teresa.ambrose@ubc.ca)

<sup>1</sup>Department of Physical Therapy, University of British Columbia, 212-2177 Wesbrook Mall, Vancouver, BC V6T1Z3, Canada

<sup>2</sup>Djavad Mowafaghian Centre for Brain Health, Vancouver Coastal Health Research Institute, University of British Columbia, Vancouver, Canada



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

multiple nights. Typically, WWA collects two types of data: 1) either triaxial or uniaxial acceleration (or motion) data; and 2) light exposure data. These data are then typically processed through a brand/software-specific scoring algorithm which is used to determine estimates of different sleep parameters. Using these scoring algorithms, WWAs are capable of providing valid and reliable estimates of *sleep duration* (total time spent sleeping), *sleep efficiency* (proportion of time spent sleeping vs. time spent in bed trying to sleep), and *wake after sleep onset* (WASO; time spent awake after sleep has been initiated and before final awakening); WWA is less accurate for determining *sleep latency* (time it takes to transition from wake to sleep) (Sadeh 2011; Kushida et al. 2001).

Scoring algorithms for determining different indices of sleep quality relies on human selection of the rest-interval of the wearer. The rest-interval is traditionally based upon manual selection of the light's out time (LO: time wearer tried to go to sleep) and got up time (GU: time the wearer got out of bed to start their day). These time points are generally based on the initial time marking a period of prolonged cessation/onset of both movement and light. Once the rest-interval has been defined, automatic sleep parameter scoring can be executed by the software program. The need to manually identify rest-intervals has led to the development of rest-interval scoring protocols to allow for reproducibility and consistency in sleep parameter estimation (Chow et al. 2016; Patel et al. 2015). These protocols typically use self-documented sleep diaries and event markers (when available) to try and improve the accuracy of sleep estimation.

The manual selection of rest-intervals is time consuming, which has led to the development of algorithms used to automatically score sleep measures of the wearer (van Hees et al. 2015, 2018; Driller et al. 2016). While light exposure data is included as a parameter in standard manual scoring, automatic algorithms have chosen to remove light as a parameter due to the perceived likelihood of the sensor being covered by the wearer's sleeve or bedding. Similar arguments can also be made toward accelerometer data, given that it is common for people to be sedentary at night before sleeping. These unavoidable aspects of actigraphy are typically why manual scoring of data is supplemented with self-report diaries and event markers. Yet, since the aim of automatic scoring is to provide minimal burden to the scorer and user as well as consistency in scoring between wearers, other methods of verification should be brought into consideration. Light exposure, while not perfect in its approximation of the rest-interval, may serve as an effective means of supplementing motion-based algorithms.

We propose a quantitative means of determining the LO and GU times using uniaxial activity and light exposure data collected from a WWA: the MotionWatch8®. The aim of this algorithm is not to improve the accuracy of sleep parameter estimation in actigraphs but to establish a consistent method of determining the rest-interval without the use of a sleep diary or event markers that is more consistent with manual rest-interval scoring protocols. The use of light as opposed to sleep diary or event marker in supplementing scoring can prove advantageous in populations where self-report compliance may be limited and inconsistent—such as in individuals experiencing cognitive decline (Fisher et al. 2006; Moul et al. 2004; Rikli 2000). As such, this algorithm will aim to create consistency from wearer-to-wearer, allowing for congruence between studies documenting sleep and for the automation of sleep analysis in further research. This study also compares two additional parameter settings – one biasing light and one biasing movement – to self-report and manual standardized protocol sleep measures. We did this in order to isolate the parameters of acceleration and illuminance data, providing insight into their usefulness in quantifying the rest-interval and their relation to one another.

## Method

We investigated the accuracy of three different automated scoring algorithms for determining the rest-interval of older adults who wore a WWA for  $\geq 14$  days of observation during the Sleep and Cognition Study (Falck et al. 2018a; Landry et al. 2015), a cross-sectional study which examined age and cognitive differences in objectively measured sleep. All participants provided written and informed consent (H14–01301).

## Participants

We recruited 154 participants from Vancouver, British Columbia. Participants were included if they met the following criteria: 1) men and women 50+ years of age living in the Metro Vancouver area; 2) scored  $> 24/30$  on the Mini-Mental State Examination (MMSE) (Kurlowicz and Wallace 1999); and 3) able to read, write, and speak English with acceptable visual and auditory acuity. We excluded individuals: 1) diagnosed with dementia of any type; 2) diagnosed with another neurodegenerative or neurological condition (e.g., Parkinson's disease or Multiple Sclerosis) which affects cognitive function and/or sleep quality; 3) taking medications which may negatively affect cognition; 4) planning to participate or currently enrolled in a clinical drug trial; or 5) unable to speak as judged by an inability to communicate by phone.

## Measures

### *Instrumentation and software*

We used the MotionWatch8® (MW8) WWA to collect activity (i.e., counts) and light data (lux) (Ancoli-Israel et al. 2015). Data were collected in 60 s epochs. Activity counts are an arbitrary unit of measurement, which is calculated for each epoch as the sum of the peak acceleration relative to a minimum acceleration threshold of 0.1 g, in a range of 0.1–8 g, sampled at a frequency of 6 Hz. The MW8 also has a lux range from 0 to 64,000 lx and samples the light exposure at a frequency of 1 Hz. The per epoch lux value recorded represents the average lux over the specified epoch length when sampled once per second. In the standardized protocol, we used the event marker time stamp button to create event marker recordings in the data to help define the rest-interval. The MW8 is the updated version of the Actiwatch7, an actigraph with evidence of validity against polysomnography in healthy adults (Mean age:  $30 \pm 6$  years; 45% female; (O'Hare et al. 2015)), and also adults with chronic insomnia (Mean age:  $41 \pm 12$  years; 78% female; (Martoni et al. 2012)). The MW8 has evidence of validity among: 1) 54 adults with suspected sleep disorders including obstructive sleep apnea, insomnia, hypersomnia, and Ehlers Danlos syndrome (Mean Age:  $53 \pm 16$  years; 61% female); and 2) 19 healthy adults (Mean Age:  $28 \pm 5$  years; 53% female; (Elbaz et al. 2012)).

After the rest-interval had been determined using one of the methods we define below (see [Rest-interval scoring methods](#)), sleep estimates were determined using the MotionWare software. The sleep/wake algorithm used by the MotionWare software is available from the manufacturer (CamNtech; [admin@camntech.co.uk](mailto:admin@camntech.co.uk)).

### *Consensus sleep diary (CSD)*

The CSD is a self-report sleep diary designed and used predominantly in insomnia research that has been shown to be effective in clinical and research settings (Carney et al. 2012). We elected to use the CSD-core which has 9-questions and is effective for both “good” and “poor” sleepers. Question's 2 and 7 were used to confirm the rest-interval defined by event marker recordings in the standardized protocol scoring method and were used to set the rest-interval in the CSD scoring method.

### *Procedure*

Each participant was asked to wear the MW8 for 14 consecutive days (Van Someren 2007). During the 14-day period, the watch was set to record uniaxial acceleration as well as light exposure. Participants were asked to press the event marker button on the watch when they got into bed and upon their final awakening in order to set the GU and LO times. Each participant was

also given the CSD and asked to fill it out each morning after they woke up. 151 (98%) of the participants had actigraphy data recorded from the specified continuous 14-day period. Twenty-one participants were excluded because either question 2 or 7 from the CSD were not complete for any single day within the 14-day study window. Of the remaining 130 participants (84.4%), 1 had irregular sleep patterns where they slept past 12:00, violating one of the programs simplifying assumptions. The remaining 129 (83.7%) participant's data were used for testing and analysis.

### *Rest-interval scoring methods*

We tested and compared five rest-interval scoring methods by looking at the sleep parameters generated from those rest-intervals using the sleep/wake scoring algorithm (Fig. 1). We highlight each of these scoring methods below.

### *CSD only method*

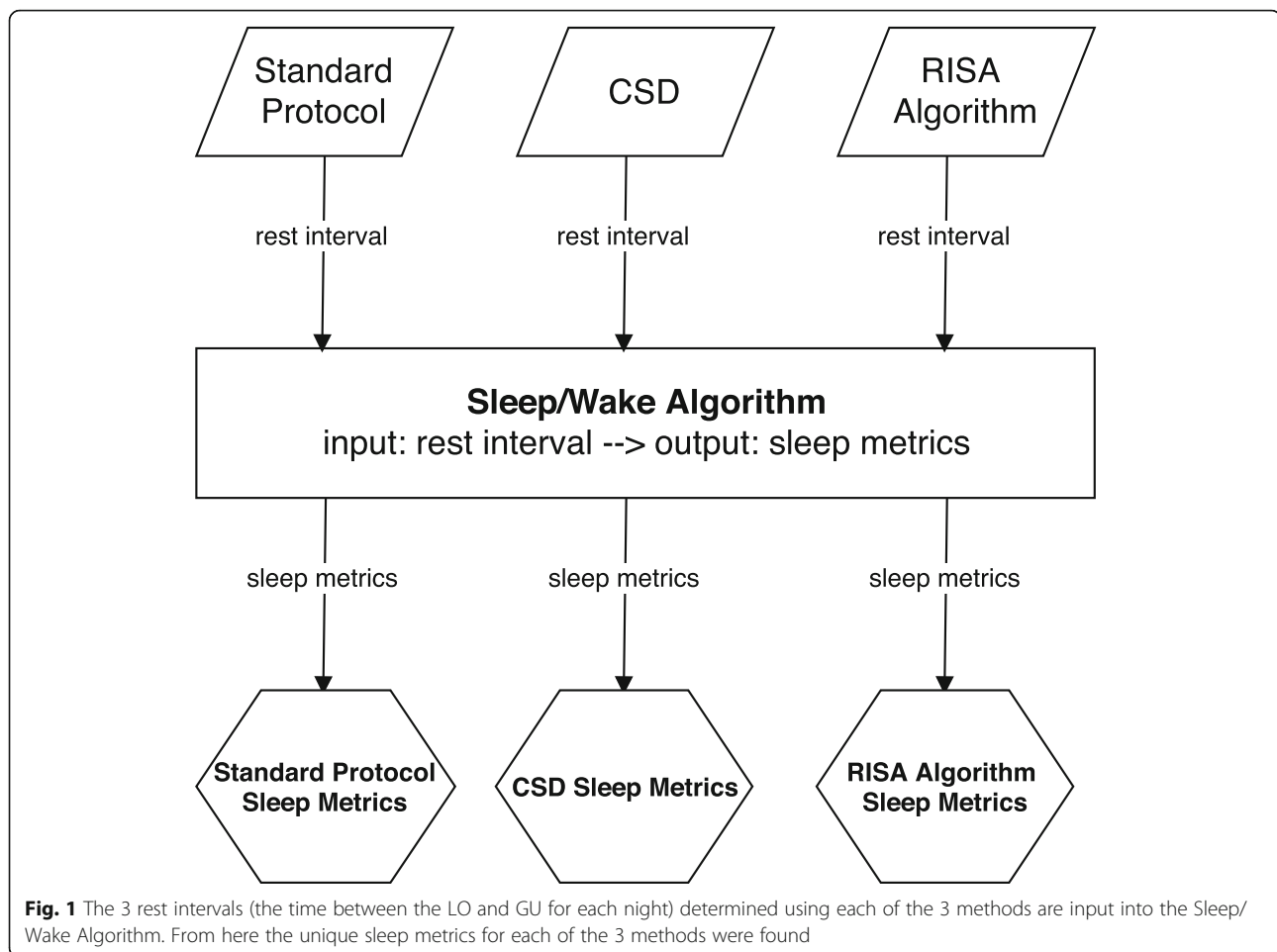
The CSD scoring method used participants answers to questions 2 and 7 of the CSD as the LO and GU time respectively to establish the rest-interval.

### *Standardized protocol*

A trained research assistant was responsible for scoring the rest-interval in the standardized protocol method. The four major sleep indices considered in order of priority were event markers, Q2 and Q7 from CSD, cessation/onset of light and cessation/onset of activity. If the event marker and CSD were within 30 min of one another, an appropriate rest-interval had been established. In cases where the indices agreement was less clear and the event marker was not within the 30 min of the CSD, the LO and GU times were set at based on the cessation/onset of light and motion. Since the precise definition of cessation/onset was not defined, whether to check light, motion or both and the number of consecutive counts of cessation/osnset was to be determined at the lab staff's discretion. When following the protocol, the data was looked at visually using the MotionWare software graphics (Fig. 2). Lab staff were asked to set the limits of the display to be 1000 and 35 for activity counts and lux respectively before proceeding with scoring. A rough interval was then selected in order to allow for accurate cursor placement. The LO and GU placements were made by adjusting a cursor displayed over top a visualization of the actigraphy data indicating the exact time (to the minute) of the cursor placement.

### *Automatic scoring algorithm*

An algorithm was created and then tested using three different threshold settings, constituting three separate scoring methods, which were then used for comparison



with the CSD and standardized protocol methods. This algorithm will be referred to as the Rest-Interval Scoring Algorithm or RISA (Fig. 3).

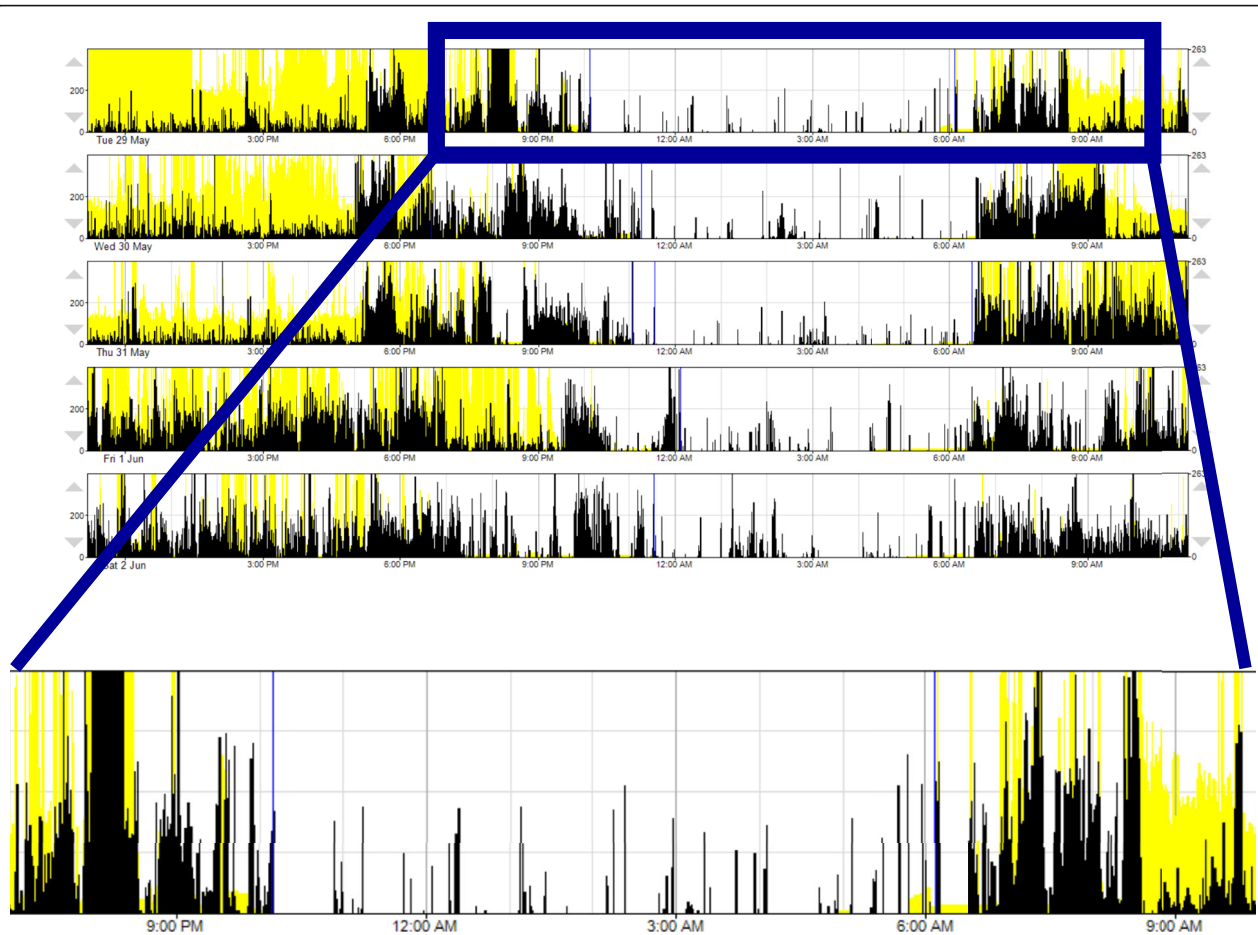
The RISA algorithm works as follows:

1. The specified period of interest (14 days in our study population) is partitioned into 24-h segments.
2. Within each 24-h segment, the window of 6 h which contains the smallest sum of epochs with an activity count below 20 and a lux count of 0 is designated as the Potential Sleep Window (PSW).
3. A 6-h window containing the potential LO time is found by segmenting data 3 h before and 3 h after the start time (ST) of the PSW.
4. A 7-h GU window is found by segmenting data 1 h before and 6 h after the end time (ET) of the PSW.
5. The potential LO and potential GU windows are then fed into two separate functions (the LO and GU functions), described below, which return the LO and GU times.

The RISA algorithm was determined heuristically (i.e. through trial and error methods). Actigraph data whose

rest-intervals had been set by the protocol method were visually inspected using the MotionWare software from an arbitrary number of participants across 4 independent studies to discern patterns in how the rest-intervals were set following the protocol method. Parameter values within the RISA and its subsequent components were initially set arbitrarily and were adjusted manually through iteration by running the algorithm on a training data set which consisted of baseline wrist-worn actigraphy recordings from the Buying Time study (Falck et al. 2018b). We elected to avoid optimization of the parameter values to prevent overfitting and allow for generalization.

The LO function abstracted away in Fig. 3 and shown in detail in Supplementary Material S1, keeps track of four counters. Each counter tracks consecutive epochs of zero lux, activity count less than 20, activity count of zero, or activity counts of zero activity and lux simultaneously. The GU function (Supplementary Material S2), keeps track of three counters. The counters track consecutive epochs of lux greater than zero, activity greater than zero, activity greater than 20, or activity and lux greater than zero simultaneously. Once a counter



**Fig. 2** Graphics used by MotionWare Software to display MW8 data recording. The yellow peaks show the lux count (a measure of light collected by the MW8), and the black peaks show the motion count

exceeds its respective threshold value, the start index is marked as the LO (or GU depending on the function) time for that day.

The threshold values within the LO function of the RISA were given 3 different sets of values; one for biasing light (BL), one for biasing motion (BM), and one which balanced both light and motion (LM). The RISA is thus the *parent algorithm* and is the underlying logical structure for LM, BM and BL. The set of *child algorithms* developed from the RISA are the LM, BM and BL. The GU function's threshold values were held constant. Each set of threshold values in the LO function constituted a separate method which was then compared to the CSD and protocol scoring methods. The BL, BM and LM threshold values are given in Supplementary Material [S3](#).

### Statistical analysis

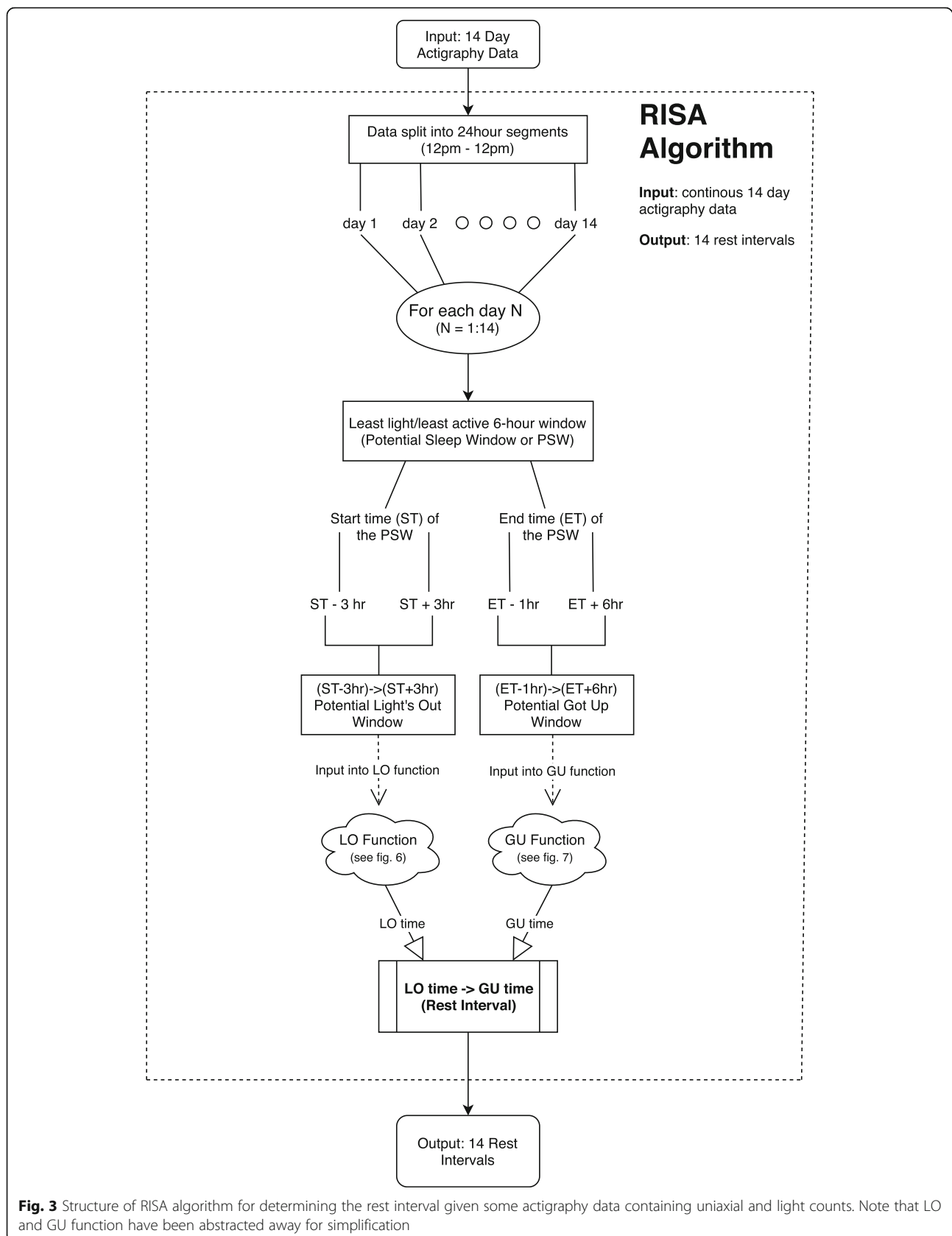
We calculated means and standard deviations for sleep duration, fragmentation, efficiency, and latency using each of the methods for estimating the rest-interval

length. We estimated between method differences in estimates of sleep quality using analysis of variance (ANOVA). Significant differences were then explored using pairwise comparisons (i.e., estimated marginal means). Further, we performed bi-variate correlation analyses to examine consistency in sleep quality estimates between methods. We subsequently conducted Bland Altman plot analyses in order to investigate differences in estimated sleep quality between the standardized protocol method and our automated methods of determining the rest-interval length.

### Results

Descriptive statistics for the study participants are given in Table 1. Most participants were female (71.3%) and retired (78.3%). The average age of participants in the study was 71.1 with a SD of 7.3. Notably, the age range of participants was relatively large, with the youngest participant being 53 and the eldest being 101.

Mean and standard deviation sleep parameter estimates according to each rest-interval estimation method



**Fig. 3** Structure of RISA algorithm for determining the rest interval given some actigraphy data containing uniaxial and light counts. Note that LO and GU function have been abstracted away for simplification



**Table 1** Descriptive statistics for study participants

Variable	Mean (SD) or N (%)
Age	71.1 (7.3)
Gender, female	92 (71.3%)
Education, >HS Diploma	92 (71.3%)
Employment, retired	101 (78.3%)
MMSE	29.0 (1.0)
MoCA	25.0 (2.6)

HS High school, MMSE Mini-Mental State Examination, MoCA Montreal Cognitive Assessment

are described in Table 2. There was not statistically significant difference between rest-interval estimation methods in sleep duration ( $F[4, 640] = 1.877$ ;  $p = 0.112$ ), fragmentation index ( $F[4, 640] = 0.398$ ;  $p = 0.810$ ) and sleep efficiency ( $F[4, 640] = 1.164$ ;  $p = 0.326$ ). There were statistically significant differences between groups when measuring sleep latency ( $F[4, 640] = 14.831$ ;  $p = 1.35 \times 10^{-11}$ ). For sleep latency, there were no significant differences between the CSD and Protocol ( $F[1256] = 0.701$ ;  $p = 0.403$ ), LM and Protocol ( $F[1256] = 0.021$ ;  $p = 0.884$ ) and LM and CSD ( $F[1, 256] = 0.427$ ;  $p = 0.514$ ) however

**Table 2** Mean, SD and  $p$ -value for each sleep metric between methods

Sleep Metric	Method	Mean	SD	$P$ -value
Fragmentation index	LM	31.0	10.4	$p > 0.05$ ( $F = 0.40$ )
	CSD	30.1	10.7	
	Protocol	30.6	10.7	
	BM	31.7	10.3	
	BL	31.1	10.4	
Sleep efficiency	LM	82.6	6.1	$p > 0.05$ ( $F = 1.16$ )
	CSD	83.4	6.2	
	Protocol	82.7	6.0	
	BM	82.6	5.7	
	BL	81.8	6.4	
Sleep duration	LM	392	48	$p > 0.05$ ( $F = 1.88$ )
	CSD	388	57	
	Protocol	404	47	
	BM	395	48	
	BL	394	48	
Sleep latency	LM	6.6	8.7	$p < 0.05$ ( $F = 14.83$ )
	CSD	5.9	8.0	
	Protocol	6.7	8.0	
	BM	3.7	4.3	
	BL	12.2	14.6	

LM Light-motion algorithm, CSD Consensus sleep diary, Protocol Standard protocol set by trained user, BM Biased-motion algorithm, BL Biased-light algorithm

**Table 3** Pearson correlations between methods for different sleep metrics

Sleep Metric	Methods	LM	CSD	Protocol	BM
Fragmentation index	CSD	0.91			
	Protocol	0.92	0.98		
	BM	0.95	0.84	0.86	
	BL	0.99	0.91	0.91	0.94
Sleep efficiency	CSD	0.88			
	Protocol	0.87	0.97		
	BM	0.96	0.88	0.89	
	BL	0.97	0.83	0.84	0.91
Sleep duration	CSD	0.91			
	Protocol	0.92	0.98		
	BM	0.95	0.84	0.86	
	BL	0.99	0.90	0.91	0.94
Sleep latency	CSD	0.20*			
	Protocol	0.18*	0.79		
	BM	0.23*	0.12**	0.09**	
	BL	0.72	0.25*	0.35	0.23**

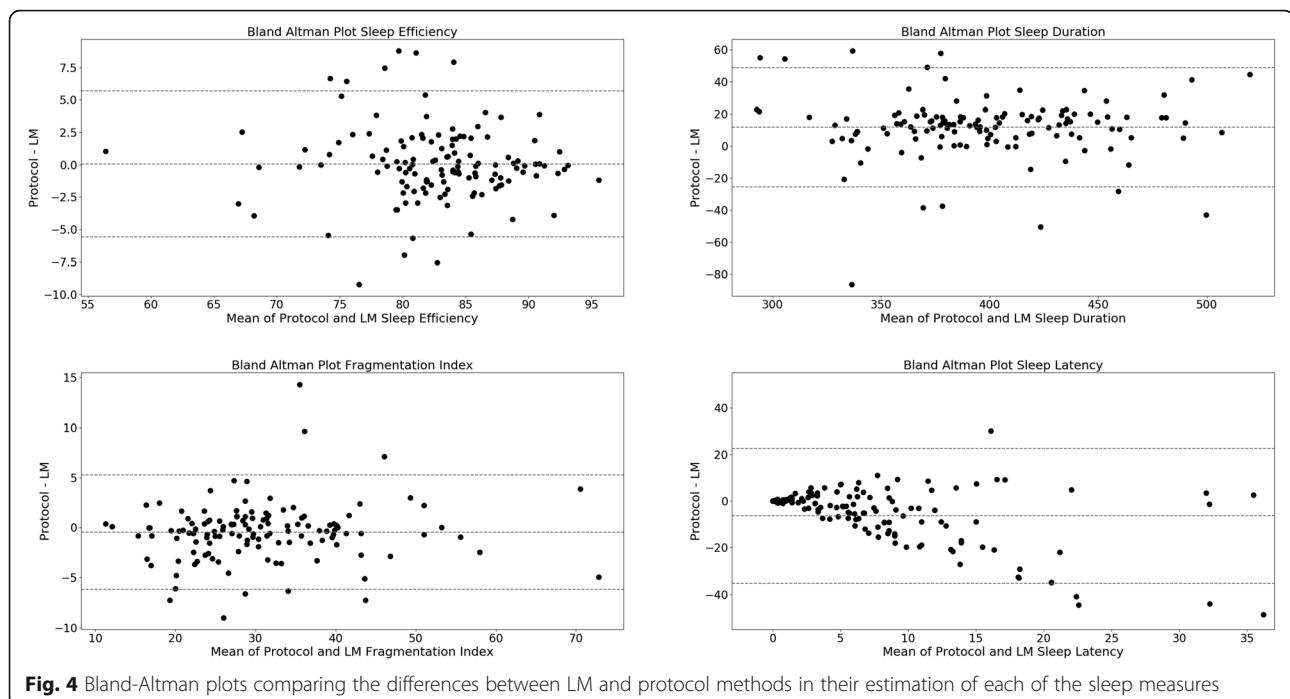
LM Light-motion algorithm, CSD Consensus sleep diary, BM Biased-motion algorithm, BL Biased-light algorithm

\* $p < 0.05$ ; \*\* $p > 0.05$

significant differences were observed between the Protocol and BM ( $F[1, 256] = 14.600$ ;  $p = 0.000167$ ).

Our bi-variate analyses are described in Table 3. The correlations observed for the metrics of sleep duration, fragmentation index and sleep efficiency all exceeded  $r = 0.80$  with statistical significance  $p < 0.001$  with the smallest correlation being between BL and CSD in their estimation of sleep efficiency ( $r = 0.83$ ,  $p < 0.001$ ). In all metrics, the strongest overall correlation to the protocol method was the CSD followed by the LM. The LM algorithm showed a comparable correlation ( $r = 0.88$ ,  $p < 0.001$ ) with the CSD as the CSD did with the protocol ( $r = 0.87$ ,  $p < 0.001$ ) for sleep efficiency. Sleep latency showed the weakest set of correlations. Of the automatic methods, the BL showed the highest correlation in sleep latency to the protocol ( $r = 0.35$ ,  $p < 0.001$ ) and CSD ( $r = 0.25$ ,  $p < 0.05$ ).

Bland-Altman plots were generated for comparing the LM and protocol scoring methods across all sleep metrics in order to observe potential differences in estimation (Fig. 4). We observed that, for sleep duration, the LM method tended to underestimate the protocol. It is worth noting that the mean difference value of roughly 15-min for sleep duration would represent a 7.5-min difference in LO and GU time placement for the LM versus protocol method. Similarly, the roughly plus-minus 30-min confidence interval represents a plus-minus 15-min difference in LO and GU placement for the LM versus protocol method. There was no observable under or over estimates



in sleep efficiency and fragmentation index, however a fanning effect was visible for sleep latency.

## Discussion

Wrist-worn actigraphy (WWA) is traditionally based on an underlying assumption: that a lack of motion is evidence of rest (Martin and Hakim 2011). Using only motion to determine the rest-interval, the certainty in estimation is limited to the extent that this assumption is true. While previous automatic scoring methods have been effective in using motion to estimate the rest-interval (van Hees et al. 2015, 2018; Driller et al. 2016), the certainty of a measurement is constrained by this limitation. To reduce these uncertainties, we have introduced the RISA algorithm, which uses light as supplemental evidence to support its estimation of the rest-interval. By reducing the uncertainty associated with the ambiguity of the data, we aim to increase the certainty in automatic rest-interval estimation, allowing for more reliable and consistent sleep metric scoring. Since WWA is a valid field-method for objectively measuring sleep (Ancoli-Israel et al. 2015), improving the reliability and consistency by which sleep quality is measured using WWA will improve our understanding about how insufficient sleep can impact older adult health.

The final algorithm selected was the LM because it is the most similar to the manual method commonly used to set the rest-interval. The LM algorithm, like the standard manual protocol method, sets the rest-interval based on the initial time marking a period of prolonged cessation of both movement and light. It is an algorithm

with predetermined logic and threshold values that are meant to emulate the same logic applied in manual scoring by trained technicians. It should be seen as a formalism of the manual scoring method, which is what gives it its advantages over manual methods where the rest-interval is set sometimes at a human's discretion. Precisely defining what a rest-interval means when looking at actigraphy data allows scoring to be consistent, and hidden uncertainties to decrease.

To assess the reliability of the RISA algorithm, it was compared against the CSD and standard protocol estimation methods across 4 different sleep metrics. The strong correlations between all methods for the metrics of sleep efficiency, sleep duration and fragmentation index provide support for the validity of the algorithm.

Using Bland-Altman plots we determined that there were no systematic differences between the LM and protocol measurements for sleep efficiency, fragmentation index and sleep duration. The underestimation of the LM when compared with the protocol for sleep duration translated to an average 7.5-min difference in LO and GU time placement, and the confidence interval a plus-minus 15-min difference. These observations are perhaps more indicative of biases of the trained technician scoring the data (Galland et al. 2014). In the case of a greater than 30-min disagreement between event-markers and self-report times, the technician placed the LO and GU times based on the cessation/onset of motion and light, the definition of which was left to the technician's discretion. Therefore, it is possible that



biases about what constitutes cessation/onset of light and motion may have caused inconsistencies in rest-interval estimation, leading to the underestimation visible in the sleep duration Bland-Altman plot.

Sleep latency showed the lowest correlations between the RISA, CSD and protocol methods. Agreement between the LM and protocol measurements for sleep latency was low, as is evident in the fanning effect visible in the Bland-Altman plot. This indicates increasingly large differences in the measurement as estimated sleep latency becomes larger. Since it has been shown that sleep latency measurements are inconsistent (Martin and Hakim 2011) (speculatively a consequence of actigraphy's underlying assumption), the differences between the two methods is to be expected.

There are several limitations to this study. Our algorithm is specific only to adults over the age of 50 years. The algorithm was designed to score actigraphy data collected from the MotionWatch8®. Since the MotionWatch8® collects motion data in arbitrary units of count and because of potential differences in the technical components of the different actigraphs, the algorithm is likely not generalizable to other tools. While we determined that LM provides comparable estimates of sleep to the standard protocol and CSD, we cannot determine whether LM provides more accurate estimates of sleep than CSD, standard protocol or WWA algorithms based only on motion. Future research will need to determine the accuracy of the RISA algorithm for estimating sleep against a gold-standard measure (e.g. polysomnography). While the rest-interval is the anchor by which sleep parameters are derived, it is not a common used metric for determining sleep quality. We therefore did not examine the accuracy of the automatic scoring method of the rest interval as compared to the manual scoring method.

Another limitation is that we did not ask participants about their natural sleeping environments, and thus we cannot account for differences in lighting between participants. For instance, we cannot determine whether participants were wearing short or long sleeves, sleeping with their arms over or under the covers, or had sleeping lights. It is important to note, however, that determining the precise sleeping environment of participants (e.g., sleeping with arms over or under the covers) would be very difficult to quantify accurately without direct observation.

Participants were asked about whether they had been diagnosed with obstructive sleep apnea. Only one participant reported an obstructive sleep apnea diagnosis; however, we did not screen for obstructive sleep apnea using the STOP-BANG questionnaire (Chung et al. 2012) or another method for determining obstructive sleep apnea risk. Untreated and undiagnosed sleep apnea could have skewed sleep data, potentially compromising

results of the algorithm. The algorithm also represents a formalism of sleep in actigraphy data developed based on human inferences from community-dwelling older adults, and is likely not generalizable to other populations. The algorithm's limitations are a consequence of these inferences. Simplifying assumptions drawn from this population - such as 12 pm–12 pm days - may not hold true for older adults with sleep disorders or irregular sleep patterns. However, it is this formalism which provides the RISA algorithm with its consistency. It is perhaps best seen as an implementation of a standardized protocol's logic, subsequently allowing for time efficient estimations and consistency across studies.

## Conclusion

In summary, the results of this study indicate that for community-dwelling older adults, the RISA algorithm provides an effective measurement of the rest-interval using data collected from the MotionWatch8® and is the first automatic algorithm to use illuminance data to supplement scoring. Future studies should address methods of formally quantifying uncertainty in automatic sleep scoring to better inform clinicians and researchers on the methods' reliability. Further investigation into the efficacy of light in other populations will also be necessary to generalize these results outside of an older adult population.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s41606-020-00051-1>.

**Additional file 1.** Flow diagram of the LO function to determine the LO times of the wearer.

**Additional file 2.** Flow diagram of the GU function to determine the GU times of the wearer.

**Additional file 3.** BL, biased-light algorithm; BM, biased-motion algorithm; LM, light-motion algorithm; IM, inactive motion; ILMM, inactive light and minimal motion; ILNM, inactive light and no motion; IL, inactive light; ALM, active light and motion; AM, active motion; AL, active light.

## Abbreviations

WWA: Wrist-worn actigraphy; LO: Lights-out; GU: Got-up; MMSE: Mini-Mental State Examination; MW8: MotionWatch8®; CSD: Consensus sleep diary; RISA: Rest-Interval Scoring Algorithm; (PSW): Potential Sleep Window; (ST): Start time; ET: End time; BL: Biasing-light algorithm; BM: Biasing-motion algorithm; LM: Balanced-light and motion algorithm; ANOVA: Analysis of variance; MoCA: Montreal Cognitive Assessment

## Acknowledgments

Not applicable.

## Authors' contributions

DB, RSF, and TLA conceived the study concept and design. DB wrote the first draft of the manuscript. RSF and TLA were responsible for initial data collection. DB and RSF conducted the statistical analyses and interpreted the results. RSF and TLA provided critical review and wrote portions of the manuscript. The author(s) read and approved the final manuscript.

### Funding

Funding was provided to TLA by the Jack Brown and Family Alzheimer's Research Foundation and the Alzheimer's Society Research Program. The funder played no part in the study design, analysis, or interpretation of results.

### Availability of data and materials

Data can be made available upon request.

### Ethics approval and consent to participate

All participants provided written and informed consent (H14–01301).

### Consent for publication

Not applicable.

### Competing interests

No competing interest to report.

Received: 30 July 2020 Accepted: 19 October 2020

Published online: 23 November 2020

### References

- Ancoli-Israel S, et al. The SBSM guide to actigraphy monitoring: clinical and research applications. *Behav Sleep Med*. 2015;13(Suppl 1):S4–S38.
- Carney CE, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep*. 2012;35(2):287–302.
- Chow CM, et al. Defining the rest interval associated with the main sleep period in actigraph scoring. *Nat Sci Sleep*. 2016;8:321–8.
- Chung F, et al. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *Br J Anaesth*. 2012;108(5):768–75.
- Driller M, McQuillan J, O'Donnell S. Inter-device reliability of an automatic-scoring actigraph for measuring sleep in healthy adults. *Sleep Sci*. 2016;9(3):198–201.
- Elbaz M, et al. Validation of a new actigraph motion watch versus polysomnography on 70 healthy and suspected sleep-disordered subjects. *J Sleep Res*. 2012;21:218.
- Falck RS, et al. The independent associations of physical activity and sleep with cognitive function in older adults. *J Alzheimers Dis*. 2018a;63(4):1469–84.
- Falck RS, et al. Buying time: a proof-of-concept randomized controlled trial to improve sleep quality and cognitive function among older adults with mild cognitive impairment. *Trials*. 2018b;19(1):445.
- Fisher SE, et al. Obtaining self-report data from cognitively impaired elders: methodological issues and clinical implications for nursing home pain assessment. *Gerontologist*. 2006;46(1):81–8.
- Galland B, et al. Challenges and emerging technologies within the field of pediatric actigraphy. *Front Psychiatry*. 2014;5:99.
- Gottlieb DJ, et al. Association of usual sleep duration with hypertension: the sleep heart health study. *Sleep*. 2006;29(8):1009–14.
- Kasasbeh E, Chi DS, Krishnaswamy G. Inflammatory aspects of sleep apnea and their cardiovascular consequences. *South Med J*. 2006;99(1):58–67 quiz 68–9, 81.
- Kurlowicz L, Wallace M. The mini-mental state examination (MMSE). *J Gerontol Nurs*. 1999;25(5):8–9.
- Kushida CA, et al. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. *Sleep Med*. 2001;2(5):389–96.
- Landry GJ, Best JR, Liu-Ambrose T. Measuring sleep quality in older adults: a comparison using subjective and objective methods. *Front Aging Neurosci*. 2015;7:166.
- Lauderdale DS, et al. Self-reported and measured sleep duration: how similar are they? *Epidemiology*. 2008;19(6):838–45.
- Marino M, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep*. 2013;36(11):1747–55.
- Martin JL, Hakim AD. Wrist actigraphy. *Chest*. 2011;139(6):1514–27.
- Martoni M, et al. Using actigraphy versus polysomnography in the clinical assessment of chronic insomnia (retrospective analysis of 27 patients). *Presse Med*. 2012;41(3):e95–e100.
- Matthews KA, et al. Similarities and differences in estimates of sleep duration by polysomnography, actigraphy, diary, and self-reported habitual sleep in a community sample. *Sleep Health*. 2018;4(1):96–103.
- Moul DE, et al. Self-report measures of insomnia in adults: rationales, choices, and needs. *Sleep Med Rev*. 2004;8(3):177–98.

- O'Hare E, et al. A comparison of radio-frequency biomotion sensors and actigraphy versus polysomnography for the assessment of sleep in normal subjects. *Sleep Breath*. 2015;19(1):91–8.
- Opp MR, Toth LA. Neural-immune interactions in the regulation of sleep. *Front Biosci*. 2003;8:d768–79.
- Patel SR, et al. Reproducibility of a standardized actigraphy scoring algorithm for sleep in a US Hispanic/Latino population. *Sleep*. 2015;38(9):1497–503.
- Rikli RE. Reliability, validity, and methodological issues in assessing physical activity in older adults. *Res Q Exerc Sport*. 2000;71(Suppl 2):89–96.
- Sadeh A. The role and validity of actigraphy in sleep medicine: an update. *Sleep Med Rev*. 2011;15(4):259–67.
- Shokri-Kojori E, et al. beta-Amyloid accumulation in the human brain after one night of sleep deprivation. *Proc Natl Acad Sci U S A*. 2018;115(17):4483–8.
- Short MA, et al. Estimating adolescent sleep patterns: parent reports versus adolescent self-report surveys, sleep diaries, and actigraphy. *Nat Sci Sleep*. 2013;5:23–6.
- van Hees VT, et al. A novel, open access method to assess sleep duration using a wrist-worn accelerometer. *PLoS One*. 2015;10(11):e0142533.
- van Hees VT, et al. Estimating sleep parameters using an accelerometer without sleep diary. *Sci Rep*. 2018;8(1):12975.
- Van Someren EJ. Improving actigraphic sleep estimates in insomnia and dementia: how many nights? *J Sleep Res*. 2007;16(3):269–75.

### Publisher's Note

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

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

