Sleep Science 9 (2016) 198–201

Contents lists available at ScienceDirect

Sleep Science

journal homepage: www.elsevier.com/locate/ssci

Full length article

Inter-device reliability of an automatic-scoring actigraph for measuring sleep in healthy adults

Matthew Driller*, Joseph McQuillan, Shannon O'Donnell

University of Waikato, Hamilton, New Zealand

ARTICLE INFO

Article history:
Received 15 June 2016
Received in revised form
2 August 2016
Accepted 11 August 2016
Available online 24 August 2016

Keywords:
Actigraphy
Polysomnography
Agreement
Validity

ABSTRACT

Actigraphy has become a common method of measuring sleep due to its non-invasive, cost-effective nature. An actigraph (Readiband™) that utilizes automatic scoring algorithms has been used in the research, but is yet to be evaluated for its inter-device reliability. A total of 77 nights of sleep data from 11 healthy adult participants was collected while participants were concomitantly wearing two Readiband™ actigraphs attached together (ACT1 and ACT2). Sleep indices including total sleep time (TST), sleep latency (SL), sleep efficiency (SE%), wake after sleep onset (WASO), total time in bed (TTB), wake episodes per night (WE), sleep onset variance (SOV) and wake variance (WV) were assessed between the two devices using mean differences, 95% levels of agreement, intraclass correlation coefficients (ICC), typical error of measurement (TEM) and coefficient of variation (CV%) analysis. There were no significant differences between devices for any of the measured sleep variables ($p > 0.05$). TST, SE, SL, TTB, SOV and WV all resulted in *very high* ICC's (>0.90), with WASO and WE resulting in *high* ICC's between devices (0.85 and 0.80, respectively). Mean differences of -2.1 and 0.2 min for TST and SL were associated with a low TEM between devices (9.5 and 3.8 min, respectively). SE resulted in a 0.3% mean difference between devices. The Readiband™ is a reliable tool for researchers using multiple devices of this brand in sleep studies to assess basic measures of sleep quality and quantity in healthy adult populations.

1. Introduction

The quantification and measurement of sleep amongst various interventional and population research studies and clinical settings is of increasing importance. Different methods of monitoring sleep have been extensively researched and validated in the literature, with little focus on the inter-device reliability of such tools. Indeed, the precision required to determine changes in sleep patterns amongst different individuals and populations is of critical importance to understanding and interpreting the results of any sleep research studies.

Although considered the gold standard method of sleep measurement, polysomnography (PSG) requires a somewhat intrusive and expensive assessment of sleep indices [1]. Moreover, PSG monitoring typically requires attendance at a sleep laboratory with specialist staff, in a foreign environment, which may be inconvenient and unnatural for most individuals. Because of this, attempts have been made to measure sleep using less-invasive methods. Such methods include sleep-logs/questionnaires and wristwatch actigraphy. The use of sleep-logs and questionnaires are common as they are in-expensive and easy to administer. However, these have been shown to have a poor relationship with objective measures of sleep [2], therefore questioning their efficacy. Wristwatch actigraphy is a non-intrusive, cost-effective tool used to estimate sleep quantity and quality which has been compared to PSG, showing accuracies of ~90% in some studies for total sleep time and sleep efficiency [3–5] and as such, are widely used in the sleep literature [3]. Actigraphy involves the use of a device housed in a wristwatch that contains a small accelerometer capable of sensing movement along any one of three axes. The accelerometer samples multiple times per second and with each limb movement, the accelerometer registers this information and stores it in an adjacent memory chip. Once the recording period has finished, the actigraph is downloaded and manually scored for sleep indices by a trained sleep technician [6]. While the process of manually scoring actigraph data has been described for its inter and intra-scorer reliability, it is difficult to make conclusions on the overall reliability of actigraphy given the variation of scoring methods, brands of actigraphs and researchers themselves [3].

Given the limitations of manually scoring actigraph files, a plethora of new actigraphy devices designed to automatically score sleep have emerged. One such device, the Readiband™ (Fatigue Science, Honolulu, USA), is gaining popularity for its use in

* Corresponding author.
E-mail address: mdriller@waikato.ac.nz (M. Driller).
Peer review under responsibility of Brazilian Association of Sleep.

http://dx.doi.org/10.1016/j.slsci.2016.08.003

sleep research studies [7–10]. The Readiband™ records data at a sample rate of 16 Hz and uses a patented algorithm to automatically score sleep data via download to the companies software. The Readiband™ has been validated against PSG, with levels of accuracy 93% being reported [11]. However, while the device has been shown to be a valid sleep measurement tool, the inter-device reliability of the Readiband™ is yet to be evaluated. Assessing the inter-device reliability for multiple devices of the same brand and model is important for researchers to have confidence that separate devices are reading in a similar and reliable manner. Indeed, this type of assessment has become commonplace in evaluating the reliability of physical activity trackers [12]. However, this type of assessment is not yet standard procedure for new actigraphs that measure sleep indices. Therefore, the purpose of the current study was to investigate the inter-device reliability of the Readiband™ by evaluating 77 nights of data from healthy adult participants concomitantly wearing two Readiband™ devices attached together.

2. Materials and methods

2.1. Participants

A total of 11 healthy adults (4 male/7 female, mean ± SD; age: 33 ± 7 years) volunteered to participate in the current study. All participants provided informed written consent before taking part in the study and were free of any diagnosed sleep disorders. Ethical approval for the study was obtained through the institutions Human Research Ethics Committee.

2.2. Methodology

Participants were required to wear two wrist actigraphs (SBV2 Readiband™, Fatigue Science, Honolulu, USA), attached together over a 7-day period to assess inter-device reliability between the two devices (ACT1 and ACT2). The Readiband™ devices have been shown to have good validity (overall accuracy of 93%) when compared to the gold standard of PSG in 50 participants undergoing overnight sleep monitoring at a sleep centre [11] and have been accepted as an approved device by the Federal Drug Administration (FDA) based on this validation. The Readiband™ has also been assessed in a mini-validation study against another actigraph (Micro Mini-Motion Loggers, Ambulatory Monitoring Inc., Ardsley, USA) [10] where the two brands of actigraph were attached together for 3 nights in 8 participants, resulting in acceptable levels of agreement for sleep duration and rest duration (r=0.84 and 0.94, respectively). In the current study, both devices were tightly secured together using electrical tape so that they could not move independently of each other and were worn on the participants' non-dominant wrist before initialization of the two devices to record data in 1-minute epochs [13]. This method of determining inter-device reliability of actigraphy monitors has been used previously [14]. Participants were required to wear the actigraph continuously for the 7-day period, with the exception of time spent in water, bathing or showering. Participants were instructed to maintain their usual sleep habits and general daily activity patterns during the monitoring period. At the conclusion of the recording period, actigraph data were wirelessly downloaded to a study computer using a Nordic 2.4 GHz ANT transceiver, which was then analyzed using Fatigue Science software (16 Hz sampling rate: Readiband™, Fatigue Science, Vancouver). The raw activity scores were translated to sleep-wake scores based on computerized scoring algorithms. The five measures obtained from the actigraphy device and software that were used as sleep indices are described in Table 1.

Table 1
Definitions of each sleep variable measures using the Fatigue Science, Readiband™ actigraph.

Sleep indices	Units	Description
Total Sleep Time (TST)	Minutes	Total time spent asleep
Sleep Efficiency (SE)	%	Total time in bed divided by total sleep time
Total Time in Bed (TTB)	Minutes	Total time spent in bed
Sleep Latency (SL)	Minutes	Time taken for sleep onset
Wake Episodes per Night (WE)	Number count	Total number of awakenings per night
Wake After Sleep Onset (WASO)	Minutes	Time spent awake after sleep onset per night
Sleep Onset Variance (SOV)	Minutes	Variation in sleep onset time
Wake Variance (WV)	Minutes	Variation in wake time
Sleep Onset Time (SOT)	Time of day (p.m.)	Time fell asleep at night
Wake Time (WT)	Time of day (a.m.)	Time woken in morning

2.3. Statistical analysis

Simple group statistics are shown as means ± standard deviations unless stated otherwise. A students paired t-test was used to compare ACT1 and ACT2 using a Statistical Package for Social Science (V. 22.0, SPSS Inc., Chicago, IL), with statistical significance set at $p \leq 0.05$. Inter-device agreements for ACT1 and ACT2 were examined using intraclass correlation coefficients (ICC) with 95% confidence intervals (95% CI) and interpreted as 0.90–1.00=*very high* correlation, 0.70–0.89=*high* correlation, 0.50–0.69=*moderate* correlation, 0.26–0.49=*low* correlation and 0.00–0.25=*little, if any* correlation [15]. The mean differences and upper and lower limits of agreement (1.96 standard deviations or 95% of a normally distributed population) between devices were determined in absolute values for TST, SL and SE. Between-device typical error of measurement (TEM) was determined using an excel spreadsheet [16] and are presented as a coefficient of variation percentage (CV%) and as absolute values. Similar to Werner et al. [17], we defined an apriori difference between the 2 devices of ≤30 min satisfactory for TST, with a difference <5% for SE satisfactory.

3. Results

There were no significant differences between devices (ACT1 and ACT2) for any of the measured sleep variables ($p > 0.05$, Table 2). There was a mean difference between devices of −2.1 ± 13.4 min over the 77 nights of data for TST. This difference was associated with a *very high* correlation and a low TEM (9.5 min) and CV (2.3%) between devices (Table 3).

Table 2
Mean ± SD values for both devices (ACT1 and ACT2) for all measured sleep variables and p-values for each comparison.

	ACT1	ACT2	P-Value
Total Sleep Time (min)	461.6 ± 86.6	459.5 ± 87.9	0.20
Sleep Efficiency (%)	83.0 ± 8.9	83.2 ± 8.9	0.73
Sleep Latency (min)	21.9 ± 20.0	21.7 ± 19.6	0.79
Total Time in Bed (min)	564.1 ± 98.7	563.2 ± 99.0	0.55
Wake After Sleep Onset (min)	11.7 ± 8.0	12.2 ± 8.4	0.32
Wake Episodes (No. per night)	3.5 ± 2.5	3.6 ± 2.6	0.72
Sleep Onset Variance (min)	−0.8 ± 74.0	−2.3 ± 75.1	0.13
Wake Variance (min)	−3.1 ± 48.4	−2.6 ± 47.3	0.37
Sleep Onset Time (time of day)	22:47 ± 0:49	22:48 ± 0:49	0.76
Wake Time (time of day)	7:03 ± 0:52	7:02 ± 0:50	0.41