HHS Public Access

Author manuscript

Physiol Behav. Author manuscript; available in PMC 2017 July 06.

Published in final edited form as:

Physiol Behav. 2016 May 01; 158: 143–149. doi:10.1016/j.physbeh.2016.03.006.

Measures of Sleep and Cardiac Functioning During Sleep Using a Multi-Sensory Commercially–Available Wristband in Adolescents:

Wearable Technology to Measure Sleep and Cardiac Functioning

Massimiliano de Zambotti[1], **Fiona C. Baker**[1,2], **Adrian R. Willoughby**[1], **Job G. Godino**[3,4], **David Wing**[3,4], **Kevin Patrick**[3,4], and **Ian M. Colrain**[1,5,*]

[1]Center for Health Sciences, SRI International, Menlo Park, CA, USA

[2]Brain Function Research Group, School of Physiology, University of the Witwatersrand, Johannesburg, South Africa

[3]Center for Wireless and Population Health Systems, Qualcomm Institute of Calit2, San Diego, USA

[4]Department of Family Medicine and Public Health, University of California, San Diego, USA

[5]Melbourne School of Psychological Sciences, University of Melbourne, Parkville, Victoria, Australia

Abstract

To validate measures of sleep and heart rate (HR) during sleep generated by a commercially-available activity tracker against those derived from polysomnography (PSG) in healthy adolescents. Sleep data were concurrently recorded using FitbitChargeHR™ and PSG, including electrocardiography (ECG), during an overnight laboratory sleep recording in 32 healthy adolescents (15 females; Age, mean±SD: 17.3±2.5 years). Sleep and HR measures were compared between FitbitChargeHR™ and PSG using paired t-tests and Bland-Altman plots. Epoch-by-epoch analysis showed that FitbitChargeHR™ had high overall accuracy (91%), high sensitivity (97%) in detecting sleep, and poor specificity (42%) in detecting wake on a min-to-min basis. On average, FitbitChargeHR™ significantly but negligibly overestimated total sleep time by 8min and sleep efficiency by 1.8%, and underestimated wake after sleep onset by 5.6min ($p<0.05$). Within FitbitChargeHR™ epochs of sleep, the average HR was 59.3±7.5 bpm, which was significantly but negligibly lower than that calculated from ECG (60.2±7.6 bpm, $p<0.001$), with no change in mean discrepancies throughout the night. FitbitChargeHR™ showed good agreement with PSG and ECG in measuring sleep and HR during sleep, supporting its use in assessing sleep and cardiac function in healthy adolescents. Further validation is needed to assess its reliability over prolonged periods of time in ecological settings and in clinical populations.

[*]*Corresponding author*. ian.colrain@sri.com (IMC), SRI International, 333 Ravenswood Avenue, Menlo Park, CA, 94025, Tel.: 650.859.3915; Fax: 650.859.2743.

Keywords

Polysomnography; Fitbit; actigraphy; wearables; adolescence; heart rate

1.1. Introduction

Adolescence is characterized by many psychophysiological developmental changes, including changes in sleep [1]. One of the most noticeable changes in sleep behavior across adolescence is an increased preference for evening activities, mostly driven by circadian factors, which leads to adolescents going to sleep later. In addition to these biological factors, social forces (e.g. early school start times) contribute to insufficient sleep and erratic sleep-wake patterns in adolescents, with high variability in sleep duration between weekdays and weekends [1]. Insufficient sleep in this population has been linked to several adverse consequences [2] including high risk for cardiovascular disease [3, 4]. This is not surprising as, the autonomic nervous system (ANS) is still developing in adolescence [5], and it is well known that sleep plays a major role in the recovery of ANS function by providing an extended period of reduced cardiovascular effort [6].

Polysomnography (PSG) is the gold standard in sleep evaluation. PSG consists of recording and integrating measures of cortical electroencephalographic activity, muscle tone, and eye movements to characterize the structure of sleep. However, because it is expensive, time consuming, and requires well-trained technicians, PSG is impractical for multiples recordings in large numbers of subjects. To avoid these limitations, researchers have used actigraphy to infer sleep/wake states, which is particularly useful in clinical evaluations, for home recordings (which are more ecologically valid), and in populations where extended recordings are important to capture a full picture of sleep variability over time (such as across adolescence) [7, 8]. Actigraphy is a non-invasive technique that uses a small accelerometer, usually placed on the non-dominant wrist, to measure an individual's movement. From these movement data, filters and algorithms have been developed to infer sleep and wake states [8]. The main advantages of actigraphy over PSG are its low cost, low intrusion, and ease of use [7]. Further, the long-lasting batteries and high-capacity storage of these devices allow for weeks of recording with little intervention from technicians after first activation. Previous research has found that, compared with PSG, actigraphy has high sensitivity (ability to detect sleep) but poor specificity (ability to detect wakefulness) [9–11]. These measures, however, vary as a function of the amount of night-time wakefulness [12] and are highly dependent on the algorithm used and the population studied [8].

Recent technological advances and, in particular, the rising popularity of wearable devices (expected to reach a value of $8.5 billion by 2020 in the United States, according to the U.S. Enterprise Wearables Market) [13] have brought to market several novel, inexpensive, and potentially useful tools to assess daily physical activity and sleep [14–22]. Technology has quickly progressed from the first accelerometer-based devices to, within only a few years, multisensory products. These devices can measure several common physiological signals in addition to movement, such as heart rate (HR), temperature, and skin conductance. It is not surprising, therefore, that researchers have started looking at these devices as an economical

way to collect data on a scale that would have proven too burdensome to collect using PSG. However, before recommending their widespread adoption as a research tool, it is necessary to evaluate the accuracy, reliability, and limitations of these tools across a variety of populations [14, 15, 21–23].

To date, only a few studies have evaluated the validity of these newer devices to detect sleep against PSG [22]. We have previously evaluated the validity of Jawbone UP™ wristband (Jawbone, San Francisco, CA) against PSG in healthy adults (N=28) [24] and in adolescents (N=65) [25]. Our results showed that, similar to standard actigraphy, the Jawbone UP™ overestimates sleep and underestimates wakefulness, and the errors become worse with increasing night-time wakefulness. Two other studies have been conducted to assess the validity of Fitbit devices (Fitbit, Inc., San Francisco, CA, USA): Classic™ (the first model released by Fitbit) and Ultra™ in measuring sleep compared with PSG. One study used a population of 24 healthy adults [26] and the other used a pediatric population of preschool and school-age children and adolescents, some of whom had obstructive sleep apnea [27]. Both studies reported high sensitivity and poor specificity for the Fitbit Classic™ and Ultra™. Also, Meltzer and colleagues [27] found greater discrepancies between PSG and Fitbit Ultra™ in adolescents (with a significant underestimation of wake after sleep onset of ~54min) than in preschool or school-age children, suggesting that adolescents may exhibit more motionless wakefulness compared to younger children.

The current study aims to evaluate the validity of sleep and HR measures during sleep collected via the FitbitChargeHR™, a promising multi-sensory wristband, against PSG in a sample of healthy adolescents.

1.2. Materials and Methods

1.2.1. Participants

Thirty-two healthy adolescents (15 females; Age, mean±SD: 17.3±2.5y; Body Mass Index, BMI, mean±SD: 22.2±3.6kg.m^{-2}; 27 Caucasian) participated in the study. They were recruited from the San Francisco Bay Area as part of an ongoing multisite study of neurodevelopment in adolescents aged 12–21 years old at baseline (the National Consortium on Alcohol and NeuroDevelopment in Adolescence, NCANDA) [28].

All participants underwent clinical screening, including the Semi-Structured Assessment for the Genetics of Alcoholism [29], to ensure they were free from major physical (e.g. heart disease, traumatic brain injury) and psychological disorders (e.g. major depression, substance use disorders). A clinical PSG also confirmed none of the participants suffered from major sleep disorders (apnea-hypopnea index and periodic limb movement index were <5 events per hour of sleep in all participants). None of the participants reported taking medication affecting sleep or cardiovascular function.

The study, including the informed consent/assent procedure, was approved by the Institutional Review Board at SRI International. Adult participants provided written consent to participate and minors provided written assent along with written consent from a parent/legal guardian. All consent and assent forms are kept on file. Participants received payment

for their participation. All data from the FitbitChargeHR™ and PSG were anonymized with the use of participant codes.

1.2.2. Procedure

After the clinical/adaptation PSG night, participants underwent another PSG night during which data from the FitbitChargeHR™ were concurrently collected.

Participants slept in sound-attenuated, temperature-controlled bedrooms in the sleep laboratory at SRI International. Lights-out and lights-on times were selected by participants. A breath alcohol test (S75 Pro, BACtrack Breathalyzers, San Francisco, CA, USA) and urine drug test (10 Panel iCup drug test kit, Instant Technologies, Inc.) were administered in the evening before each PSG night to ensure that none of the participants had recently used alcohol or drugs.

1.2.3. Polysomnography

A standard PSG recording (6 electroencephalographic [EEG] leads: F3/4, C3/4, O1/2 referenced to the contralateral mastoids; submental electromyogram [EMG]; bilateral electrooculogram [EOG]; electrocardiogram [ECG]) was performed according to the American Academy of Sleep Medicine (AASM) rules [30]. EEG, EOG and EMG were sampled at 256Hz. EEG and EOG were band-pass filtered at 0.3–35Hz and EMG was filtered at 10–100Hz. The ECG signal was collected at 512Hz using Ag/AgCl Meditrace surface spot electrodes in a modified lead II Einthoven configuration, and pass-band filtered at 0.3–70Hz. R-waves were automatically detected by a customized MatLab (MathWorks, Natick, MA, USA) algorithm and visually inspected; normal-to-normal interbeat-intervals (ms) calculated and heart rate derived (HR, bpm).

Sleep was scored (Wake, N1, N2, N3 and REM) in 30s epochs according to AASM criteria [30] and standard sleep parameters were calculated including: time in bed (TIB, min), sleep onset latency as the first epoch of any sleep stage (SOL, min), REM latency (REML, min), wake after sleep onset (WASO, min), total sleep time (TST, min), sleep efficiency (SE, TST/TIB*100, %), percentage of time spent in each stage of sleep calculated as a percentage of TST (N1, N2, N3 and REM, %), and total number of awakenings (awakening index) and arousals (arousal index) per hour of sleep.

During the clinical/adaptation night, bilateral anterior tibialis EMG (to assess leg movements), thoracic and abdominal piezoelectric bands (to measure respiratory effort), airflow using nasal cannula and thermistor, and oxygen saturation (pulse oximeter) were also collected and used to confirm the absence of sleep disorders.

1.2.4. FitbitChargeHR™

The FitbitChargeHR™ (Fitbit Inc., San Francisco, CA, USA) is a commercially available wireless activity-tracking wristband powered by the MEMS tri-axial accelerometer. Detailed information about the Fitbit device can be found on the Fitbit help website (http://help.fitbit.com/). Briefly, the FitbitChargeHR™ generates measures of activity (e.g. calories burned, minutes of activity at different intensities, steps) and sleep (e.g. time to fall asleep,

overall amount of sleep at night). In addition, the FitbitChargeHR™ measures heart rate (HR) continuously by detecting changes in blood volume using PurePulse™ LED lights.

For the sleep recording, we used the "*Normal*" setting, indicated as "*appropriate for most users*" by the Fitbit help documentation (Sleep tracking FAQs). HR tracking was set to "*Auto*". The FitbitChargeHR™ band was placed on the wrist of the participant's non-dominant hand, "*a finger's width above the wrist bone" and "not too tight*" as suggested by Fitbit documentation (Heart rate FAQs). Sleep mode was activated by pressing and holding the FitbitChargeHR™ button for more than 2s both in the evening and in the morning in synchronization with PSG lights-off and lights-on by the sleep lab technicians.

We extracted the following sleep measures from the Fitbit dashboard: "*you were in bed for*" (equivalent to PSG TIB), "*time to fall asleep*" (equivalent to PSG SOL, min) and the "*actual sleep time*" (equivalent to PSG TST). WASO was calculated by subtracting "*actual sleep time*" and "*time to fall asleep*" from "*you were in bed for*". Sleep efficiency was derived from the Fitbit by dividing "*actual sleep time*" by "*you were in bed for*" and multiplying by 100 to correspond with the SE calculation using PSG measures.

To allow for min-by-min comparison of FitbitChargeHR™ and PSG measures, we also linked our FitbitChargeHR™ devices to Fitabase (Small Steps Labs LLC), a third-party research platform designed to collect data from the Fitbit (https://www.fitabase.com). Through Fitabase we were able to download minute-by-minute measures of HR (bpm) and sleep state – categorized as Sleep, Restless and Wake states (Restless and Wake states were combined and are considered Wake in the analyses) – for each night. Fitbit dashboard measures of time in bed, time to fall asleep, and actual sleep time matched those calculated by summing epochs of wake/sleep as obtained by Fitabase.

1.2.5. Statistical Analyses

Two participants were excluded from analyses of sleep measures but not HR due to their extremely poor sleep with PSG measures >3SD from the mean. Three participants were excluded from HR analyses due to an unknown failure (probably due to the position of the wrist-band) in the FitbitChargeHR™ recording of HR.

The PSG sleep measures and their equivalents derived from the FitbitChargeHR™ were compared using paired t-tests. The agreement between PSG and FitbitChargeHR™ sleep measures was estimated using the Bland-Altman method [31]. The mean difference (or *Bias*) between FitbitChargeHR™ and PSG sleep measures was calculated along with the standard deviation, 95% CI, and lower and upper agreement limits (mean difference ±1.96SD). A positive *Bias* indicates that the FitbitChargeHR™ underestimates PSG sleep measures and a negative *Bias* indicates that the FitbitChargeHR™ overestimates them. The number of participants falling outside the Bland-Altman agreement limits and the *a priori* satisfactory ranges (a difference between PSG and FitbitChargeHR™ ≤30min for TST and WASO and ≤5% for SE) [25–27, 32, 33] is provided.

Epoch-by-Epoch (EBE) analysis of PSG and FitbitChargeHR™ sleep/wake states was also performed. Min-by-min FitbitChargeHR™ "Wake" and "Sleep" epochs were obtained

through the Fitabase research platform. To match the min-by-min resolution, PSG epochs were re-coded as 'Sleep' if both 30s epochs were scored as N1, N2, N3 or REM. They were re-coded as "Wake" if one or both of the PSG 30s epochs were scored as "Wake" [34]. FitbitChargeHR™ and PSG epochs were compared using accuracy (PSG epochs identified correctly as "Sleep" and "Wake" by FitbitChargeHR™ divided by the overall number of epochs), sensitivity (proportion of PSG epochs identified correctly as "Sleep" by FitbitChargeHR™) and specificity (proportion of PSG epochs identified correctly as "Wake" by FitbitChargeHR™). We also computed two additional indices of performance – the predicted value for sleep (PVS, proportion of FitbitChargeHR™ epochs identified correctly as "Sleep" by PSG) and the predicted value for wakefulness (PVW, proportion of FitbitChargeHR™ epochs identified correctly as "Wake" by PSG) [7].

The comparison between FitbitChargeHR™ and ECG measures of HR was performed on a min-by-min basis using FitbitChargeHR™ "Sleep" epochs only (Fig 1).

Mean HR during FitbitChargeHR™ sleep epochs derived from the ECG and FitbitChargeHR™ were compared using a paired t-test. A Bland-Altman plot is provided to show the agreement between HR generated by ECG and the FitbitChargeHR™. In order to evaluate potential changes across the night in FitbitChargeHR™ performance for HR assessment, repeated measures ANOVAs were run on the mean difference in HR (ECG minus FitbitChargeHR™) using *Time of night* (1^{st}, 2^{nd}, 3^{rd}, 4^{th}, 5^{th} and 6^{th} sextiles) as the within-subject factor.

Normality was tested using the Shapiro-Wilk W test and $p<0.05$ was considered significant in all the analyses.

1.3. Results

1.3.1. PSG and FitbitChargeHR™ Sleep Measures

PSG and FitbitChargeHR™ sleep measures are provided in Table 1. Bland-Altman plots for TST, SOL, WASO and SE are provided in Fig 2; *Bias*, SD, ±95% CI, and upper and lower agreement limits are provided in Table 2.

TST ($t=2.09$, $p=0.045$), WASO ($t=-2.13$, $p=0.041$) and SE ($t=2.21$, $p=0.035$) derived from FitbitChargeHR™ significantly differed from those derived from PSG. However, 86.7% of the sample for TST and SE, and 90% of the sample for WASO fell within *a priori* set clinically satisfactory ranges. Four of the 30 participants exceeding these limits for TST (difference >30min) and SE (difference >5%), and three of 30 participants exceeding these limits for WASO (difference >30min).

On average, FitbitChargeHR™ overestimated TST by 8min and underestimated WASO by 5.6min with three participants falling outside the Bland-Altman agreement limits. FitbitChargeHR™ overestimated SE by 1.8% with two participants falling outside the agreement limits. From the Bland-Altman plots, it is noticeable that PSG-FitbitChargeHR™ discrepancies are lower in those participants having relatively good PSG sleep with greater discrepancies emerging when PSG sleep is more disrupted.

1.3.2. Epoch-by-Epoch Analysis

Overall, the FitbitChargeHR™ showed high accuracy in detecting sleep/wake state (90.9±4.7%; 95%CI: 89.1–92.7%), with high sensitivity (97.1±2.3%; 95%CI: 96.2–97.9%) for detecting sleep and high predictive value in detecting sleep (92.9±4.7%; 95%CI: 91.1–94.6), but lower specificity for detecting wake on a min-by-min basis (42.4±15.9%; 95%CI: 36.5–48.3%) and predictive value in detecting wake (65.0±18.0%; 95%CI: 58.3–71.7%).

1.3.3. Comparison between Electrocardiogram and Fitbit PurePulse Heart Rate Monitoring

Mean FitbitChargeHR™ HR (59.3±7.5bpm) was significantly, but negligibly, lower than ECG HR (60.2±7.6bpm; $p<0.001$). On average, Bland-Altman plots show that the FitbitChargeHR™ underestimated ECG HR by 0.88bpm (±0.43bpm) with two participants falling outside the agreement limits (upper, 1.72bpm and lower, 0.04bpm agreement limits) (Fig 3).

There was no significant *Time of night* effect of for agreement between FitbitChargeHR™ and ECG HR measures ($F_{1,28}=1.21$, $p=0.307$).

1.4. Discussion

This study evaluated the validity of the FitbitChargeHR™ in measuring sleep and HR during sleep in healthy adolescents. FitbitChargeHR™ sleep and HR measures were compared against gold standard techniques to evaluate sleep and HR, i.e. PSG and ECG respectively. To our knowledge, this is the first study examining the validity of the FitbitChargeHR™ in detecting sleep and HR during sleep. Overall, our data indicated that in healthy adolescents, the FitbitChargeHR™ provides acceptably accurate measures of sleep and heart function during sleep.

The FitbitChargeHR™ performed well in detecting overall PSG-equivalent sleep measures in healthy adolescents with 86.7%–90% of the sample falling within *a priori* set clinically satisfactory minimum differences from PSG (≤30min for TST and WASO and ≤5% for SE). Also, Bland-Altman analyses showed that measures of dispersion of FitbitChargeHR™-PSG differences (such as the ±95%CI of Biases) were quite narrow for all sleep measures with high discrepancies between measurements in some participants with more disrupted PSG sleep. Similarly, we have shown that another commercial fitness-tracker device (Jawbone UP™) has reduced validity against PSG in individuals with high amounts of WASO [24, 25].

Indices derived from epoch-by-epoch analysis indicated that while the sensitivity of FitbitChargeHR™ to detect sleep was high (97%), the specificity (wake detection) was poor (42%). However, both of these values may have been influenced by limitations in methodology necessary to match the one minute resolution of the Fitbit device; PSG one-minute epochs are coded as "wake" even if only one PSG 30s epoch is scored as "Wake" (see section 1.2.5). This raises the possibility that a PSG epoch could have been misclassified as “wake” if a single wake period of greater than 15 seconds occurred in one of the two combined 30 second periods. However, low specificity (compared to an overall high accuracy in detecting sleep) is a common finding for devices that use detection of motion to

score sleep-wake epochs, as not all wakefulness is associated with movement [8]. Several other methodological challenges need to be considered in the interpretation of the sensitivity and specificity indices. The period considered for the analysis was restricted to the night-time sleep period and thus relatively small periods of wakefulness (<6% of the time in bed) are available for analysis. High sensitivity could thus be obtained by classifying all epochs as sleep. It is also possible that reliability of human PSG scoring (where a 10% error is considered acceptable) may bias calculations of specificity of actigraphic-based devices compared against PSG [11, 35]. From a user point of view, it is interesting to note that measures of "What percentage of time the FitbitChargeHR™ accurately detected sleep (Predictive value for sleep, PVS) and wake (Predictive value for wake, PVW) as true PSG sleep and wake", indicated that FitbitChargeHR™ is accurate 93% of the time when it scores an epoch as sleep and 65% of the time when it scores an epoch as wake.

The results from the FitbitChargeHR™ appear to be more promising than those reported from two earlier models of Fitbit devices. In studies using the FitbitClassic™ in adults [26] and the FitbitUltra™ in pediatric populations [27], the devices largely overestimated/underestimated SE (>14%), TST (>60min) and WASO (>50min). In both studies, min-by-min sleep/wake epochs were manually extracted from Fitbit output graphs and epoch-by-epoch analysis was performed. Earlier Fitbit devices showed high sensitivity and poor specificity (98% and 20% respectively, in adults [26] and 87% and 52% respectively, in the pediatric population [27]. Different actigraphs and scoring algorithms, and even different models of the same actigraph used across studies, may contribute to discrepant findings between studies. No information is available on how different Fitbit models (e.g. FitbitClassic™, FitbitUltra™, FitbitChargeHR™) detect sleep/wake states, and thus we can only speculate about potential intra-model differences. It is important to highlight that unlike earlier Fitbit models, the FitbitChargeHR™ was designed to be worn exclusively on the wrist. The placement of actigraphs on the body is a well-known contributor to differences in the measurement of movement, and this design choice may have influenced the device's measurement capabilities. A major difference between our study and earlier studies is the sleep efficiency of the population studied. In the studies of Montgomery-Downs et al. [26], and Meltzer et al. [27], average PSG SE was 79.5% [26] and 83.4%, respectively [27], while in our study average PSG SE was 90.8%. Performance of actigraphs of any type in assessing sleep is limited in populations with low sleep efficiency or in cases where there is a high amount of motionless wake, which the actigraph scores as sleep [8, 12]. Additionally, our group recently showed a strong linear association between age and the discrepancies between PSG and another commercially available fitness-tracker in assessing sleep in a group of healthy adolescents [25]. Thus, age-related differences in PSG architecture need to also be considered when interpreting results from different studies. Also, the potential misclassification of awakenings across the night (as in part reflected by a low specificity of the device) could be a concern in clinical populations (e.g. sleep breathing-related disorders, insomnia disorder). Further studies need to assess the reliability of FitbitChargeHR™ in capturing the pattern of wake across multiple nights and when sleep is highly disrupted.

One of the main advantages of the FitbitChargeHR™ is its ability to measure HR concurrently with movement. The detection of HR from the FitbitChargeHR™ during sleep was accurate compared with ECG HR, and the small differences in minute-to-minute HR

were constant across the night. We did not assess the validity of the FitbitChargeHR™ during the wake period given the low amount of wakefulness exhibited by our sample and the predictive value of wake index which indicated that on average, 35% of the time, wake epochs detected by the FitbitChargeHR™ were not in agreement with PSG based wake epochs.

The FitbitChargeHR™ uses an unknown proprietary algorithm, and it remains unclear whether or not the algorithm combines movement and HR data to improve the accuracy of sleep/wake state detection after sleep onset has occurred. Also, HR and HR variability measures fluctuate across the night under circadian and sleep-dependent influences [36]. The accurate detection of these fluctuations, particularly the measurement of heart rate variability (not supported by the current Fitbit models), may be highly informative in efforts to accurately assess variation in nighttime patterns of vagal activity, an important indication of potential health risk [37, 38].

Knowing the limits and the advantages of these novel activity-trackers informs researchers and clinicians about the circumstances in which these wearables can be validly used. Previous research using older Fitbit models (Flex™ and One™) over multiple nights showed high rates of missed sleep data [39, 40]. However, in our study, the FitbitChargeHR™ reliably detected sleep in all participants, albeit on one night in a controlled laboratory environment. On the other hand, we had 3 participants in which the FitbitChargeHR™ failed to correctly detect HR, which seemed to be due to the incorrect positioning of the wristband. This factor needs to be taken into account when multiple nights of recordings are involved. Several other issues need to be further investigated including the changes in FitbitChargeHR™ performance with changes in sleep/wake patterns across age as well as in relation to motion during sleep and wakefulness; the reliability of FitbitChargeHR™ when sleep is extremely poor (such as in clinical populations); and the validity of FitbitChargeHR™ in different populations.

While the investigation of commercial devices is only just beginning, it would be beneficial to have more transparent details about Fitbit device function and more flexibility in accessing raw data (currently available only from third parties, e.g. Fitabase).

In conclusion, we have shown that the FitbitChargeHR™ provides a valid measure of sleep and heart rate during sleep in healthy adolescents with typical sleep patterns. These relatively inexpensive, easy-to-use, and techno-attractive devices have the potential for paving the way to track sleep in large adolescent populations across multiple days to investigate relationships between sleep and other behaviors (e.g. physical activity, substance use).

Conclusions

Sleep profoundly changes across adolescence and has implications for health, safety, performance, substance use/abuse and mood. The explosion of wearable health technology provides new opportunities to measure sleep metrics over prolonged periods alone, and in conjunction with observation of other behaviors. This study showed the validity of a

promising and popular activity tracker's assessment of sleep and heart rate during sleep in adolescents.

Acknowledgments

We would like to thank Lena Kardos, David Sugarbaker, David Dresser, Stephanie Claudatos, and Sarah Inkelis for their effort in the data collection process. We would also like to thank Fitabase (Small Steps Labs LLC, San Diego, CA, USA) for allowing the use of their platform. We would like to emphasize that this was an independent investigation and none of the authors have a relationship with Fitbit Inc. (San Francisco, CA, USA).

References

1. Colrain I, Baker F. Changes in sleep as a function of adolescent development. Neuropsychol Rev. 2011; 21:5–21. [PubMed: 21225346]
2. McKnight-Eily L, Eaton D, Lowry R, Croft J, Presley-Cantrell L, Perry G. Relationships between hours of sleep and health-risk behaviors in US adolescent students. Prev Med. 2011; 53:271–273. [PubMed: 21843548]
3. Javaheri S, Storfer-Isser A, Rosen CL, Redline S. Sleep Quality and Elevated Blood Pressure in Adolescents. Circulation. 2008; 118:1034–1040. [PubMed: 18711015]
4. Narang I, Manlhiot C, Davies-Shaw J, Gibson D, Chahal N, Stearne K, et al. Sleep disturbance and cardiovascular risk in adolescents. CMAJ. 2012; 184:E913–E920. [PubMed: 23027917]
5. Silvetti M, Drago F, Ragonese P. Heart rate variability in healthy children and adolescents is partially related to age and gender. Int J Cardiol. 2001; 81:169–174. [PubMed: 11744133]
6. Trinder J, Waloszek J, Woods M, Jordan A. Sleep and cardiovascular regulation. Pflugers Arch. 2012; 463:161–168. [PubMed: 22038322]
7. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak C. The role of actigraphy in the study of sleep and circadian rhythms. Sleep. 2003; 26:342–392. [PubMed: 12749557]
8. Van de Water A, Holmes A, Hurley D. Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography-a systematic review. J Sleep Res. 2011; 20:183–200. [PubMed: 20374444]
9. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. Sleep Med Rev. 2011; 15:259–267. [PubMed: 21237680]
10. Martin J, Hakim A. Wrist actigraphy. Chest. 2011; 139:1514–1527. [PubMed: 21652563]
11. de Souza L, Benedito-Silva A, Pires M, Poyares D, Tufik S, Calil H. Further validation of actigraphy for sleep studies. Sleep. 2003; 26:81–85. [PubMed: 12627737]
12. Paquet J, Kawinska A, Carrier J. Wake detection capacity of actigraphy during sleep. Sleep. 2007; 30:1362–1369. [PubMed: 17969470]
13. Compass Intelligence. US Enterprise Wearables Market: 5-Year Forecast, 2014–2019. 2015
14. Grifantini K. How's My Sleep?: Personal sleep trackers are gaining in popularity, but their accuracy is still open to debate. Pulse, IEEE. 2014; 5:14–18.
15. Fischer S. Sleep on It: Sleep might just be the most important part of daily health? and the biggest new target for biomedical engineering. Pulse, IEEE. 2014; 5:8–13.
16. Chen K, Bassett D Jr. The technology of accelerometry-based activity monitors: current and future. Med Sci Sports Exerc. 2005; 37:S490–S500. [PubMed: 16294112]
17. Steins D, Dawes H, Esser P, Collett J. Wearable accelerometry-based technology capable of assessing functional activities in neurological populations in community settings: a systematic review. J Neuroeng Rehabil. 2014; 11:36. [PubMed: 24625308]
18. Guo F, Li Y, Kankanhalli MS, Brown MS. An evaluation of wearable activity monitoring devices. Proceedings of the 1st ACM international workshop on Personal data meets distributed multimedia: ACM. 2013:31–34.
19. Dobkin B. Wearable motion sensors to continuously measure real-world physical activities. Curr Opin Neurol. 2013; 26:602–608. [PubMed: 24136126]

20. Lowe S, Olaighin G. Monitoring human health behaviour in one's living environment: A technological review. Med Eng Phys. 2014; 36:147–168. [PubMed: 24388101]
21. Lee J, Finkelstein J. Consumer sleep tracking devices: a critical review. Stud Health Technol Inform. 2015; 210:458–460. [PubMed: 25991187]
22. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act. 2015; 12:159. [PubMed: 26684758]
23. Poyares D, Hirotsu C, Tufik S. Fitness Tracker to Assess Sleep: Beyond the Market. Sleep. 2015; 38:1351–1352. [PubMed: 26237767]
24. de Zambotti M, Claudatos S, Inkelis S, Colrain I, Baker F. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chrobiol Int. 2015; 37:1024–1028.
25. de Zambotti M, Baker F, Colrain I. Validation of sleep-tracking technology compared with polysomnography in adolescents. Sleep. 2015; 38:1461–1468. [PubMed: 26158896]
26. Montgomery-Downs H, Insana S, Bond J. Movement toward a novel activity monitoring device. Sleep Breath. 2012; 16:913–917. [PubMed: 21971963]
27. Meltzer LJ, Hiruma L, Avis K, Montgomery-Downs HE, Valentin J. Comparison of a Commercial Accelerometer with Polysomnography and Actigraphy in Children and Adolescents. Sleep. 2015; 38:1323–1330. [PubMed: 26118555]
28. Brown S, Brumback T, Tomlinson K, Cummins K, Thompson W, Nagel B, et al. The National Consortium on Alcohol and NeuroDevelopment in Adolescence (NCANDA): Characterizing risk and resilience for alcohol use in adolescents. Journal of Studies on Alcohol and Drugs. J Stud Alcohol Drugs. under review.
29. Bucholz K, Cadoret R, Cloninger C, Dinwiddie S, Hesselbrock V, Nurnberger J Jr, et al. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. J Stud Alcohol. 1994; 55:149–158. [PubMed: 8189735]
30. Iber, C., Ancoli-Israel, S., Chesson, A., Quan, SF. for the American Academy of Sleep Medicine. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. 1st. Westchester, Illinois: American Academy of Sleep Medicine: American Academy of Sleep Medicine; 2007.
31. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1:307–310. [PubMed: 2868172]
32. Meltzer L, Walsh C, Traylor J, Westin A. Direct comparison of two new actigraphs and polysomnography in children and adolescents. Sleep. 2012; 35:159–166. [PubMed: 22215930]
33. Werner H, Molinari L, Guyer C, Jenni O. Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns. Arch Pediatr Adolesc Med. 2008; 162:350–358. [PubMed: 18391144]
34. Sadeh A, Sharkey K, Carskadon M. Activity-based sleep-wake identification: an empirical test of methodological issues. Sleep. 1994; 17:201–207. [PubMed: 7939118]
35. Tryon W. Issues of validity in actigraphic sleep assessment. Sleep. 2004; 27:158–165. [PubMed: 14998254]
36. Trinder J, Kleiman J, Carrington M, Smith S, Breen S, Tan N, et al. Autonomic activity during human sleep as a function of time and sleep stage. J Sleep Res. 2001; 10:253–264. [PubMed: 11903855]
37. Chakko S, Mulingtapang R, Huikuri H, Kessler K, Materson B, Myerburg R. Alterations in heart rate variability and its circadian rhythm in hypertensive patients with left ventricular hypertrophy free of coronary artery disease. Am Heart J. 1993; 126:1364–1372. [PubMed: 8249794]
38. Mølgaard H, Sørensen K, Bjerregaard P. Circadian variation and influence of risk factors on heart rate variability in healthy subjects. Am J Cardiol. 1991; 68:777–784. [PubMed: 1892086]
39. Lillehei AS, Halcón LL, Savik K, Reis R. Effect of inhaled lavender and sleep hygiene on self-reported sleep issues: a randomized controlled trial. J Altern Complement Med. 2015; 21:430–438. [PubMed: 26133206]
40. Baroni A, Bruzzese J-M, Di Bartolo CA, Shatkin JP. Fitbit Flex: an unreliable device for longitudinal sleep measures in a non-clinical population. Sleep Breath. in press.

Highlights

- Accuracy of FitbitChargeHR™ to assess sleep and heart rate during sleep is tested
- Fitbit device is reliable in detecting standard polysomnographic (PSG) metrics
- Fitbit device performed well in detecting heart rate during sleep
- Similar to standard actigraphy, Fitbit device had lower ability in detecting wake
- FitbitChargeHR™ may be a valid alternative to PSG in healthy populations

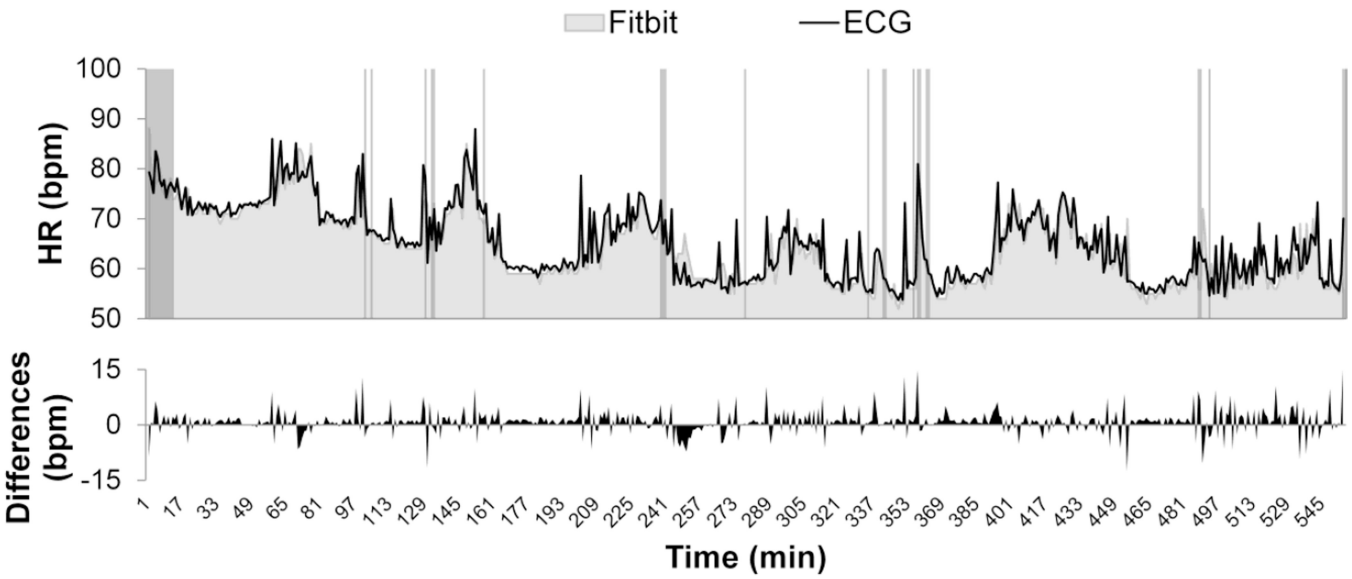

Fig 1.

Minute-by-minute heart rate (HR) across the night for a representative participant. Upper panel: HR (bpm) derived from the electrocardiogram (ECG, black) is superimposed on FitbitChargeHR™ HR (grey). Vertical bars show periods of wakefulness detected by FitbitChargeHR™ and are excluded from the analyses. Lower panel: Differences in HR (ECG minus FitbitChargeHR™) across the night.

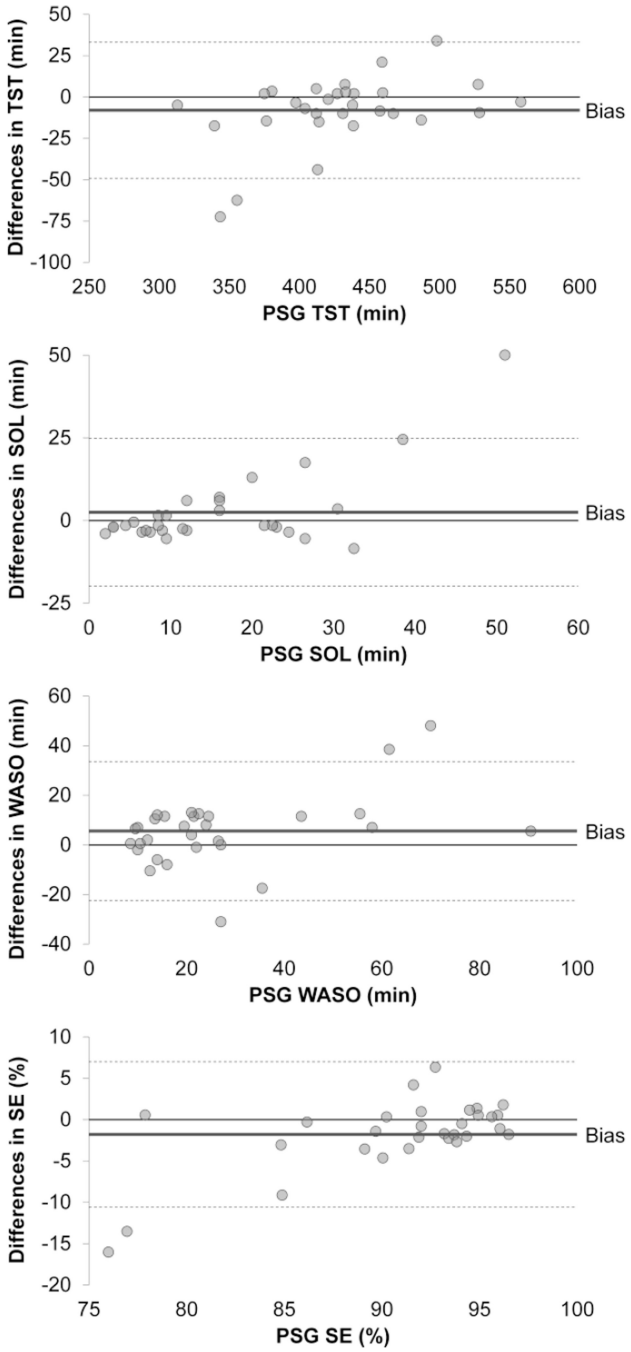

Fig 2.

Bland-Altman plots for polysomnographic (PSG) and FitbitChargeHR™ total sleep time (TST), sleep onset latency (SOL), wake after sleep onset (WASO) and sleep efficiency (SE) in a sample of 30 healthy adolescents. Bias (thick solid line), upper and lower agreement limits (mean difference ±1.96SD, thin dotted lines) are shown. The PSG measure is shown on the x-axis and the difference (PSG minus FitbitChargeHR™) is shown on the y-axis.

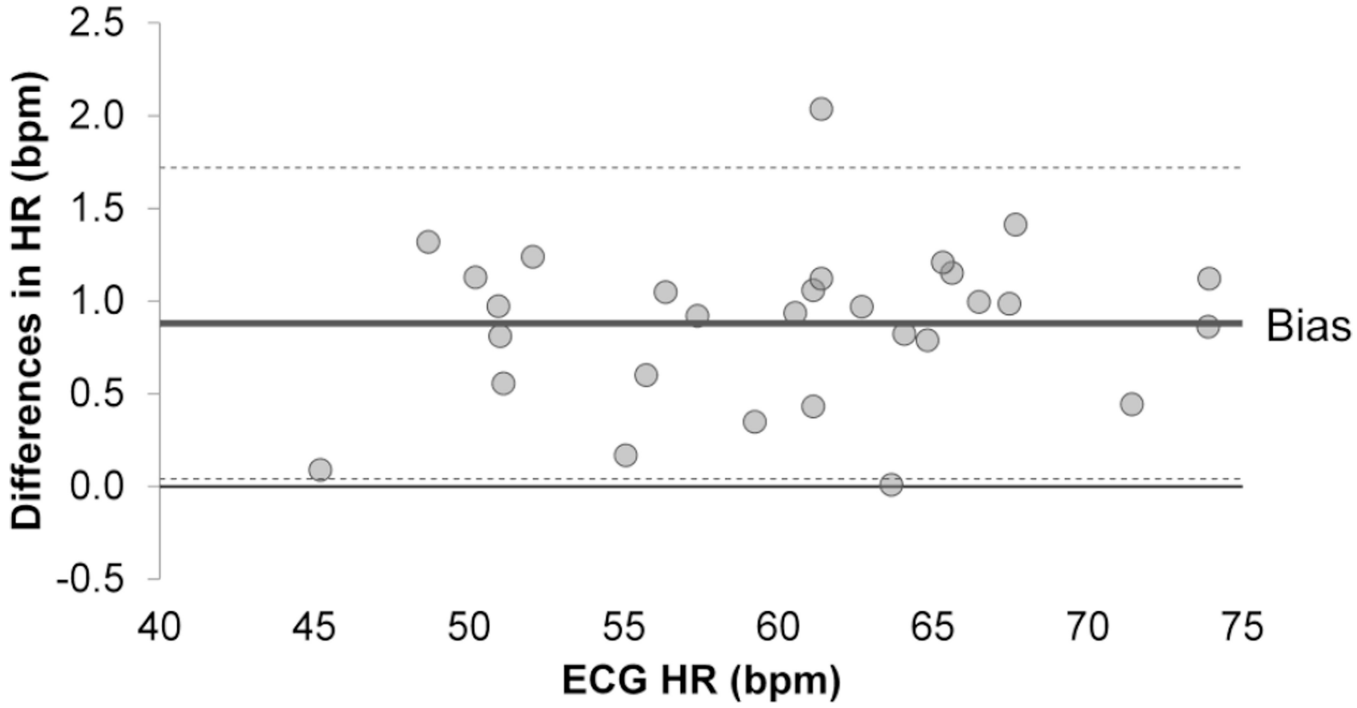

Fig 3.

Bland-Altman plot for electrocardiographic (ECG) and FitbitChargeHR™ heart rate (HR) in 29 adolescents. Bias (ECG minus Fitbit HR, thick solid line) and upper and lower agreement limits (mean difference ±1.96SD, thin dotted lines) are shown.

Table 1

Mean, SD, ±95%CI for polysomnographic and FitbitChargeHR™ sleep outcomes from one overnight recording in 30 healthy adolescents. **REM**, rapid eyes movement; **SE**, sleep efficiency; **SOL**, sleep onset latency; **TIB**, time in bed; **TST**, total sleep time; **WASO**, wake after sleep onset.

	Polysomnography			**FitbitChargeHR™**		
	Mean ± SD	**−95%CI**	**+95%CI**	**Mean ± SD**	**−95%CI**	**+95%CI**
Lights-off (hh:mm)	23:13 ± 1:03	22:50	23:37			
Lights-on (hh:mm)	07:04 ± 00:45	06:47	07:21			
TIB (min)*	471.3 ± 55.9	450.4	492.2			
TST (min)	427.9 ± 56.8	406.7	449.1	435.9 ± 51.1	416.8	455.0
SOL (min)	16.2 ± 11.7	11.8	20.5	13.6 ± 9.4	10.1	17.1
WASO (min)	27.2 ± 20.4	19.6	34.9	21.7 ± 19.3	14.5	28.9
SE (%)	90.8 ± 5.6	88.7	92.9	92.3 ± 4.2	91.0	94.2
REM latency (min)	116.2 ± 38.1	102.0	130.5			
Awakening Index (No./h sleep)	3.1 ± 1.2	2.6	3.5			
Arousal Index (No./h sleep)	8.3 ± 2.9	7.3	9.4			
Time in N1 (%TST)	5.4 ± 2.1	4.6	6.2			
Time in N2 (%TST)	49.1 ± 6.6	46.6	51.5			
Time in N3 (%TST)	23.7 ± 7.3	21.0	26.4			
Time in REM (%TST)	21.8 ± 4.7	20.1	23.6			

*Time in bed was equivalent between PSG and FitbitChargeHR™ by design

Table 2

Bias, SD, ±95%CI of Bias, upper and lower agreement limits for Bland-Altman plots between FitbitChargeHR™ and polysomnographic sleep outcomes in 30 adolescents. **SE**, sleep efficiency; **SOL**, sleep onset latency; **TST**, total sleep time; **WASO**, wake after sleep onset.

	Bias ± SD	−95%CI of Bias	+95%CI of Bias	Lower agreement limit	Upper agreement limit
TST (min)	−8.0 ± 21.0	−15.9	−0.2	−49.2	33.1
SOL (min)	2.5 ± 11.4	−1.7	6.8	−19.8	24.9
WASO (min)	5.6 ± 14.3	0.2	10.9	−22.4	33.6
SE (%)	−1.8 ± 4.5	−3.4	−0.1	−10.6	7.0