MEGHNA P. MANSUKHANI, MD
Center for Sleep Medicine, Mayo Clinic, Rochester, MN

BHANU PRAKASH KOLLA, MD, MRCPsych
Center for Sleep Medicine and Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN

Apps and fitness trackers that measure sleep: Are they useful?

ABSTRACT

Consumers have become increasingly interested in using fitness trackers and smartphone applications to quantify sleep. The devices claim to measure various sleep parameters, with the result that patients are now showing the data from their devices to their clinicians with concerns about their quantity or quality of sleep. In general, these devices have major shortcomings and limited utility, as they have not been thoroughly evaluated in clinical populations.

KEY POINTS

Wearable fitness trackers tend to perform better than smartphone applications, which are more prone to interference from bed partners and pets.

Sleep data from tracking devices are less reliable in patients with fragmented sleep and insomnia.

In normal sleepers, devices tend to measure sleep duration with reasonable accuracy, so that one can tell if a patient is getting too little sleep or reassure someone who is getting enough sleep.

Devices may help identify patients with poor sleep hygiene or atypical circadian rhythms.

Dr. Mansukhani has disclosed research funding from ResMed.

doi:10.3949/ccjm.84a.15173

More and more consumers are using wearable devices and smartphones to monitor and measure various body functions, including sleep. Many patients now present their providers with sleep data obtained from their phones and other devices. But can these devices provide valid, useful clinical information?

This article describes common sleep tracking devices available to consumers and the mechanisms the devices probably use to distinguish sleep from wakefulness (their algorithms are secret), the studies evaluating the validity of device manufacturers' claims, and their clinical utility and limitations.

DEVICES ARE COMMON

Close to 1 in 10 adults over age 18 owns an activity tracker, and sales are projected to reach \$50 billion by 2018.[1] Even more impressive, close to 69% of Americans own a smartphone,[1] and more than half use it as an alarm clock.[2]

At the same time that these devices have become so popular, sleep medicine has come of age, and experts have been pushing to improve people's sleep and increase awareness of sleep disorders.[3,4] While the technology has significantly advanced, adoption of data from these devices for clinical evaluation has been limited. Studies examining the validity of these devices have only recently been conducted, and companies that make the devices have not been forthcoming with details of the specific algorithms they use to tell if the patient is asleep or awake or what stage of sleep the patient is in.

WHAT ARE THESE DEVICES?

Consumer tracking devices that claim to measure sleep are easily available for purchase and

TABLE 1

Fitness tracking devices vs conventional methods of evaluating sleep

Device	Comparator	N	Study population	Results
Fitbit[13]	Polysomnography and actigraphy	24	Healthy adults with no history of symptoms of sleep disorders Mean age 26.1	Sensitivity for sleep 97.8% Specificity for wakefulness 19.8% Overestimated total sleep time and sleep efficiency and underestimated wake time after sleep onset
Fitbit Ultra[14]	Polysomnography and actigraphy	63	Children and adolescents undergoing overnight clinical polysomnography Mean age 9.7	'Normal' mode overestimated total sleep time and sleep efficiency and underestimated wake time after sleep onset 'Sensitive' mode underestimated total sleep time and sleep efficiency and overestimated wake time after sleep onset
Jawbone UP[15]	Polysomnography	65	Healthy adolescents Mean age 15.8	Overestimated total sleep time and sleep efficiency and underestimated wake time after sleep onset, no difference in sleep onset latency No clear correlation between 'light' and 'deep' sleep and conventional polysomnographic sleep stages
Jawbone UP[16]	Polysomnography	28	Midlife women Mean age 50.1	Sensitivity for sleep 96% Specificity for wakefulness 37% Overestimated total sleep time and sleep onset latency and underestimated wake time after sleep onset
Jawbone UP[17]	Polysomnography and actigraphy	64	Children and adolescents with suspected sleep-disordered breathing Mean age 8.4	Sensitivity for sleep 92% Specificity for wakefulness 66% No difference from polysomnography on total sleep time, sleep efficiency, sleep onset latency, and wake time after sleep onset Compared with actinography, overestimated sleep onset latency and underestimated wake time after sleep onset

include wearable fitness trackers such as Fitbit, Jawbone UP, and Nike+ Fuelband. Other sleep tracking devices are catalogued by Ko et al.[5] Various smartphone applications (apps) are also available.

Fitness trackers, usually worn as a wrist band, are primarily designed to measure movement and activity, but manufacturers now claim the trackers can also measure sleep. Collected data are available for the user to review the following day. In most cases, these trackers display sleep and wake times; others also claim to record sound sleep, light sleep, and the number and duration of awakenings. Most fitness trackers have complementary apps available for download that graphically display the data on smartphones and interact with social media to allow users to post their sleep and activity data.

More than 500 sleep-related apps are available for download to smartphones in the iTunes app store[5]; the Sleep Cycle alarm clock app was among the top 5 sleep-tracking apps downloaded in 2014.[6] Because sleep data collection relies on the smartphone being placed on the user's mattress, movements of bed partners, pets, and bedding may interfere with results. In most cases, the apps display data in a format similar to that of fitness trackers. Some claim to determine the optimal sleep phase for the alarm to wake the user.

HOW DOES THE TECHNOLOGY WORK?

Older activity-tracking devices used single-channel electroencephalographic recordings or multiple physiologic channels such as galvanic skin response, skin temperature, and

TABLE 2

Smartphone sleep-tracking applications vs conventional methods of evaluating sleep

App	Comparator	N	Study population	Results
Sleep Time[18]	Polysomnography	20	Volunteers with no sleep disorders Mean age 39.5	Sensitivity for sleep 89.9% Specificity for wakefulness 50% No clear correlation between Sleep Time and polysomnography for sleep onset latency and sleep stages Did not awaken subjects exclusively from light sleep
Motion X 24/7[17]	Polysomnography and actigraphy	64	Children, adolescents with suspected obstructive sleep apnea Mean age 8.4	Overestimated total sleep time and sleep efficiency Underestimated sleep onset latency and wake time after sleep onset
Toss 'N' Turn[19]	Pittsburgh Sleep Quality Index (sleep diary)	27	Volunteers Mean age 34	Reasonable accuracy for classifying good and poor sleepers per diary No clear correlation between app and diary for bedtime, wake time, and sleep duration

heat flux to measure activity to determine transitions between periods of sleep and wakefulness.[7,8]

None of the currently available consumer sleep tracking devices discloses the exact mechanisms used to measure sleep and wakefulness, but most appear to rely on 3-axis accelerometers,[9] ie, microelectromechanical devices that measure front-to-back, side-to-side, and up-and-down motion and convert the data into an activity count. Activity counts are acquired over 30- or 60-second intervals and are entered into algorithms that determine if the pattern indicates that the patient is awake or asleep. This is the same method that actigraphy uses to evaluate sleep, but most actigraphs used in medicine disclose their mechanisms and provide clinicians with the option of using various validated algorithms to classify the activity counts into sleep or awake periods.[9–11]

ARE THE MEASURES VALID?

Only a few studies have examined the validity and accuracy of current fitness trackers and apps for measuring sleep.

The available studies are difficult to compare; most have been small and used different actigraphy devices for comparison. Some tested healthy volunteers, others included people with suspected or confirmed sleep disorders, and some had both types of participants. In many studies, the device was compared with polysomnography for only 1 night, making the "first-night effect" likely to be a confounding factor, as people tend to sleep worse during the first night of testing. Technical failures for the devices were noted in some studies.[12] In addition, some currently used apps may use different platforms than the devices used in these studies, limiting the extrapolation of results.

Two fitness tracking devices (Fitbit[13,14] and Jawbone UP[15–17]) were compared with polysomnography and actigraphy in several studies in children and adults **(Table 1)**. The devices tended to overestimate sleeping time, sleep efficiency, and latency to sleep onset and underestimate awake time after falling asleep. Some studies noted that differences were more pronounced for those with the most disturbed sleep.

Close to 1 in 10 adults owns an activity tracker

As with fitness trackers, few studies have been done to examine the validity of smartphone apps.[5] Findings of 3 studies are summarized in **Table 2**.[17–19] In addition to tracking the duration and depth of sleep, some apps purport to detect snoring, sleep apnea, and periodic limb movements of sleep. Discussion of these apps is beyond the scope of this review.

In general, sleep tracking devices are fair to good at detecting sleep but poor at determining wakefulness. They are inaccurate for determining absolute sleep parameters (ie, total sleep time, sleep efficiency, wake time after sleep onset, and sleep onset latency) and in distinguishing the different sleep stages compared to polysomnography. Age-related differences have been found between consumer sleep devices compared with polysomnography and actigraphy-derived measures; because adults are likelier to lie still when awake, activity monitors are prone to overestimate sleep time in adults. Comparisons in patients with sleep apnea are conflicting.[12,15] Claims that the "sensitive" mode may be appropriate for users with sleep disorders are thus far unsubstantiated.

ARE THE DEVICES CLINICALLY USEFUL?

Although a thorough history remains the cornerstone of a good evaluation of sleep problems, testing is sometimes essential, and in certain situations, objective data can complement the history and clarify the diagnosis.

Polysomnography remains the gold standard for telling when the patient is asleep vs awake, diagnosing sleep-disordered breathing, detecting periodic limb movements and parasomnias, and aiding in the diagnosis of narcolepsy.

Companies have not been forthcoming about how the devices measure sleep

Actigraphy, which uses technology similar to fitness trackers, can help distinguish sleep from wakefulness, reveal erratic sleep schedules, and help diagnose circadian rhythm sleep disorders. In patients with insomnia, actigraphy can help determine daily sleep patterns and response to treatment.[20] It can be especially useful for patients who cannot provide a clear history, eg, children and those with developmental disabilities or cognitive dysfunction.

Consumer sleep tracking devices, like actigraphy, are portable and unobtrusive, providing a way to measure sleep duration and demonstrate sleep patterns in a patient's natural environment. Being more accessible, cheaper, and less time-consuming than clinical tests, the commercially available devices could be clinically useful in some situations, eg, for monitoring overall sleep patterns to look for circadian sleep-wake disorders, commonly seen in shift workers (shift work disorder) or adolescents (delayed sleep-wake phase disorder); or in patients with poor motivation to maintain a sleep diary. Because of their poor performance in clinical trials, they should not be relied upon to distinguish sleep from wakefulness, quantify the amount of sleep, determine sleep stages, and awaken patients exclusively from light sleep.

Discerning poor sleep hygiene from insomnia

Patients with insomnia tend to take longer to go to sleep (have longer sleep latencies), wake up more (have more disturbed sleep with increased awakenings), and have shorter sleep times with reduced sleep efficiencies.[21] Sleep tracking devices tend to be less accurate for patients with short sleep duration and disturbed sleep, limiting their usefulness in this group. Furthermore, patients with insomnia tend to underestimate their sleep time and overestimate sleep latency; some devices also tend to overestimate the time to fall asleep, reinforcing this common error made by patients.[22,23]

On the other hand, data from sleep tracking devices could help distinguish poor sleep hygiene from an insomnia disorder. For example, the data may indicate that a patient has poor sleep habits, such as taking long daytime naps or having significantly variable time in and out of bed from day to day. The total times asleep and awake in the middle of the night may also be substantially different on each night, which would also possibly indicate poor sleep hygiene.

Detecting circadian rhythms

A device may show that a patient has a clear circadian preference that is not in line with his or her daily routines, suggesting an underlying circadian rhythm sleep-wake disorder. This may be evident by bedtimes and wake times that are consistent but substantially out of sync with one's social or occupational needs.

Measuring overall sleep duration

In people with normal sleep, fitness trackers perform reasonably well for measuring overall sleep duration. This information could be used to assess a patient with daytime sleepiness and fatigue to evaluate insufficient sleep as an etiologic factor.

Table 3 summarizes how to evaluate the data from sleep apps and fitness tracking de-

TABLE 3

Assessing data from sleep apps and fitness trackers

If the data from a device or app indicates…	The patient may have…
Substantially different daily bedtimes and wake times Significant variations in total sleep time Long naps Long awake periods in the middle of the night	**Poor sleep hygiene**
Substantial difference from what patient reports	**Misperception of sleep**
Consistent delay or early time to bed	**Circadian rhythm sleep-wake disorder**

vices for clinical use. While these features of consumer sleep tracking devices could conceivably help in the above clinical scenarios, further validation of devices in clinical populations is necessary before their use can be recommended without reservation.

ADVISING PATIENTS

Patients sometimes present to clinicians with concerns about the duration of sleep time and time spent in various sleep stages as delineated by their sleep tracking devices. Currently, these devices do not appear to be able to adequately distinguish various sleep stages, and in many users, they can substantially underestimate or overestimate sleep parameters such as time taken to fall asleep or duration of awakenings in the middle of the night. Patients can be reassured about this lack of evidence and should be advised to not place too much weight on such data alone.

Sleep "goals" set by many devices have not been scientifically validated. People without sleep problems should be discouraged from making substantial changes to their routines to accommodate sleep targets set by the devices. Patients should be counseled about the pitfalls of the data and can be reassured that little evidence suggests that time spent in various sleep stages correlates with adverse daytime consequences or with poor health outcomes.

Some of the apps used as alarm clocks claim to be able to tell what stage of sleep people are in and wait to awaken them until they are in a light stage, which is less jarring than being awakened from a deep stage, but the evidence for this is unclear. In the one study that tested this claim, the app did not awaken participants from light sleep more often than is likely to occur by chance.[17] The utility of these apps as personalized alarm clocks is still extremely limited, and patients should be counseled to obtain an adequate amount of sleep rather than rely on devices to awaken them during specific sleep stages.

Dozens of sleep-tracking apps are available for download to smartphones

The rates for discontinuing the use of these devices are high, which could limit their utility. Some surveys have shown that close to 50% of users stop using fitness trackers; 33% stop using them within 6 months of obtaining the device.[24] Also, there is little evidence that close monitoring of sleep results in behavior changes or improved sleep duration. Conversely, the potential harms of excessive monitoring of one's sleep are currently unknown. ■

REFERENCES

1. **Rock Health**. The future of biosensing wearables. http://rockhealth.com/reports/the-future-of-biosensing-wearables/. Accessed March 16, 2017.
2. **Time, Inc.** Your wireless life: results of Time's mobility poll. http://content.time.com/time/interactive/0,31813,2122187,00.html. Accessed March 16, 2017.
3. **Office of Disease Prevention and Health Promotion (ODPHP)**. Healthy people 2020. Sleep health. www.healthypeople.gov/2020/topics-objectives/topic/sleep-health. Accessed March 16, 2017.
4. **Consensus Conference Panel; Watson NF, Badr MS, Belenky G, et al**. Joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society on the recommended amount of

sleep for a healthy adult: methodology and discussion. J Clin Sleep Med 2015; 11:931–952.
5. **Ko PT, Kientz JA, Choe EK, Kay M, Landis CA, Watson NF**. Consumer sleep technologies: a review of the landscape. J Clin Sleep Med 2015; 11:1455–1461.
6. **Investor Place Media, LLC**. Top iTunes picks: Apple names best apps of 2014. http://investorplace.com/2014/12/apple-best-apps-of-2014-aapl/#.VIYeE9LF98E/. Accessed April 13, 2017.
7. **Sunseri M, Liden CB, Farringdon J, et al**. The SenseWear armband as a sleep detection device. Internal publication.
8. **Shambroom JR, Fábregas SE, Johnstone J**. Validation of an automated wireless system to monitor sleep in healthy adults. J Sleep Res 2012; 21:221–230.
9. **John D, Freedson P**. ActiGraph and Actical physical activity monitors: a peek under the hood. Med Sci Sports Exerc 2012; 44(suppl 1):S86–S89.
10. **Sadeh A, Sharkey KM, Carskadon MA**. Activity-based sleep-wake identification: an empirical test of methodological issues. Sleep 1994; 17:201–207.
11. **Kripke DF, Hahn EK, Grizas AP, et al**. Wrist actigraphic scoring for sleep laboratory patients: algorithm development. J Sleep Res 2010; 19:612–619.
12. **Meltzer LJ, Marcus CL**. Reply: caffeine therapy for apnea of prematurity: long-term effect on sleep by actigraphy and polysomnography. Am J Respir Crit Care Med 2014; 190:1457–1458.
13. **Montgomery-Downs HE, Insana SP, Bond JA**. Movement toward a novel activity monitoring device. Sleep Breath 2012; 16:913–917.
14. **Meltzer LJ, Hiruma LS, Avis K, Montgomery-Downs H, Valentin J**. Comparison of a commercial accelerometer with polysomnography and actigraphy in children and adolescents. Sleep 2015; 38:1323–1330.
15. **de Zambotti M, Baker FC, Colrain IM**. Validation of sleep-tracking technology compared with polysomnography in adolescents. Sleep 2015; 38:1461–1468.
16. **de Zambotti M, Claudatos S, Inkelis S, Colrain IM, Baker FC**. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chronobiol Int 2015; 32:1024–1028.
17. **Toon E, Davey MJ, Hollis SL, Nixon GM, Horne RS, Biggs SN**. Comparison of commercial wrist-based and smartphone accelerometers, actigraphy, and PSG in a clinical cohort of children and adolescents. J Clin Sleep Med 2016; 12:343–350.
18. **Bhat S, Ferraris A, Gupta D, et al**. Is there a clinical role for smartphone sleep apps? Comparison of sleep cycle detection by a smartphone application to polysomnography. J Clin Sleep Med 2015; 11:709–715.
19. **Min JK, Doryab A, Wiese J, Amini S, Zimmerman J, Hong JI**. Toss 'n' turn: smartphone as sleep and sleep quality detector. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Toronto, Ontario, Canada: ACM; 2014:477-486.
20. **Morgenthaler T, Alessi C, Friedman L, et al; Standards of Practice Committee; American Academy of Sleep Medicine**. Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. Sleep 2007; 30:519-529.
21. **Lichstein KL, Durrence HH, Taylor DJ, Bush AJ, Riedel BW**. Quantitative criteria for insomnia. Behav Res Ther 2003; 41:427–445.
22. **Carskadon MA, Dement WC, Mitler MM, Guilleminault C, Zarcone VP, Spiegel R**. Self-reports versus sleep laboratory findings in 122 drug-free subjects with complaints of chronic insomnia. Am J Psychiatry 1976; 133:1382–1388.
23. **Perlis ML, Giles DE, Mendelson WB, Bootzin RR, Wyatt JK**. Psychophysiological insomnia: the behavioural model and a neurocognitive perspective. J Sleep Res 1997; 6:179–188.
24. **Endeavour Partners LLC**. Inside wearables: how the science of human behavior change offers the secret to long-term engagement. http://endeavourpartners.net/assets/Wearables-and-the-Science-of-Human-Behavior-Change-EP4.pdf. Accessed March 16, 2017.

***ADDRESS**: Meghna P. Mansukhani, MD, Center for Sleep Medicine, Mayo Clinic, 2nd Street SW, Rochester, MN 55905; Mansukhani.Meghna@mayo.edu*