EDITORIAL

Fitness Tracker to Assess Sleep: Beyond the Market

Commentary on de Zambotti et al. Validation of sleep-tracking technology compared with polysomnography in adolescents. SLEEP 2015;38: 1461–1468.

Dalva Poyares, MD, PhD; Camila Hirotsu, PhD; Sergio Tufik, MD, PhD

Department of Psychobiology, Universidade Federal de São Paulo (UNIFESP), Brazil

In this issue of *SLEEP*, de Zambotti and colleagues[1] make an important contribution to a topic that has been the subject of a great deal of recent discussion in this field: the lack of validation of new technologies for sleep monitoring. In the light of the increasing availability of micro electromechanical sensor systems,[2] we have witnessed an explosion of mobile applications and self-monitoring devices.[3] Indeed, several apps and small sensors that are able to self-monitor behavior and health parameters have become popular, particularly in smartphones. Most sleep activity devices, especially those based on accelerometers (a technology that uses lack of movement to determine sleep) fail to reflect the complexity of the sleep state, although they do enable consumers to manage to some extent their health data with the purpose of promoting general well-being. However, the companies that make these products do not claim to be selling a "medical device." To date, there has been no evidence of any positive effects on health from using these devices; but the question is how aware of this is the average consumer?

Another factor that has encouraged the development of health self-management devices is the increase in chronic diseases and aging populations. It is possible that these devices may reach a level of complexity that will enable patients to better manage their chronic diseases. However, before this happens, the technologies will need to be properly validated in both healthy and chronic disease patients. Regarding accelerometer-based sleep-wake trackers, whether these devices reliably estimate fragmented sleep, or the sleep of people with different sleep disorders is a matter in need of scientific study. There have been some studies showing that even the most well-validated actigraphy systems used by among sleep researchers still have problems with specificity when estimating sleep parameters in general or clinical populations.[4,5]

A recent review identified a critical absence of supporting evidence for the advertised functions and benefits in the majority of the devices.[2] Six out of seven devices did not provide any information about sensor accuracy and output validity. The main issue, however, is that although there is little evidence concerning the benefits of electronic devices that track sleep, activity, diet, and health, they are being widely used. Given their growing use, it is prudent to not only assess the efficacy and sensitivity of these devices, but also the type of user. For example, it is difficult to estimate the cognitive effect of controlling sleep/wake on insomnia patients, in particular if the sleep tracker output is unreliable.

Some studies have evaluated the accuracy of these apps when measuring activity or calories used. However, the study by de Zambotti is original because it evaluated the accuracy of a popular self-tracker device, Jawbone UP, compared with polysomnography (PSG), in adolescents, a population in which technology and self-trackers are very popular.[1] According to the results, Jawbone UP overestimated total sleep time (TST) and underestimated wake after sleep onset (WASO); however, the differences were not clinically meaningful. When the Bland Altman plot was analyzed, there was good agreement between PSG and Jawbone UP in TST and WASO, corroborating previous studies.[4,6] On the other hand, a recent study by Bhat et al. showed that a well-known sleep monitoring app (Azumio Inc.) failed to correlate with PSG results and performed poorly in reporting absolute sleep parameters or sleep stages.[7]

Interestingly, in their sample of adolescents, de Zambotti demonstrated that there was a progression from underestimation to overestimation of PSG total sleep time with Jawbone UP with advancing age.[1] This result suggests that the accuracy of devices based on accelerometers may decrease with age to the point that the authors suggest that age is an important factor to be considered in the development of novel algorithms for motion analysis in sleep/wake assessment.

In a previous paper by de Zambotti, the authors also evaluated the agreement between Jawbone UP and PSG assessing sleep in a sample of 28 midlife women.[8] For standard actigraphy, Jawbone UP had high sensitivity in detecting sleep and low specificity in detecting wake. Also, the sleep tracker overestimated TST and sleep onset latency, and underestimated WASO, with greater discrepancies in nights with more disrupted sleep compared to PSG. The higher discrepancy found in nights with more disrupted sleep is a warning not to use the devices with insomnia patients, a condition that manifests itself in disrupted sleep and is prevalent among midlife women. This type of limitation may be partially overcome by using extended data recording, to take into account daily variability and provide data on circadian rhythm.[8]

Another output found by Zambotti in Jawbone UP device is "sound sleep" and "light sleep."[1] The former was significantly associated with N2 and REM sleep, but not with N3 (assumed to be the deepest sleep stage). Also, the author found that "light sleep" was positively associated with the PSG arousal index, awakening index and time in N2 and N3. The N1 stage,

Submitted for publication July, 2015
Accepted for publication July, 2015
Address correspondence to: Dalva Poyares, MD, PhD, Department of Psychobiology, Universidade Federal de São Paulo (UNIFESP), Rua Napoleao de Barros 925, São Paulo, 04024-002, Brazil; Email: dpoyares@gmail.com or poyares@unifesp.br

which could be considered "light sleep," did not appear to be significant in the models. One may wonder whether this inaccuracy could be higher in patients with sleep disorders, namely sleep disordered breathing, insomnia, or periodic leg movements.

Kim and colleagues evaluated in two studies the general profile of female college students who were supposedly familiar with apps. They used self-trackers for activity, sleep, and diet monitoring for at least 90 consecutive days. The authors reported that what they called perceived threat, perceived usefulness, and perceived ease of use significantly affected the health consumer's attitude and behavioral intention.[9,10] They also identified subtle behavioral and emotional changes caused by self-tracking, self-reflection, self-management, and data recording. They noticed examples that included failing to meet self-set goals when users temporarily felt relieved from the constant "survey" by the self-tracker, and altering their daily behaviors in order to simplify the recording process. It is possible that the results would have been different if the study had been carried out in patients with sleep disorders, but we still do not know what the effect of constant monitoring would be, in particular in the sleep/wake and self-attitude of insomnia patients.

In the light of current literature and with the characteristics of available sleep-trackers, we find that their use cannot be recommended for older patients and patients with medical conditions, especially without professional medical supervision. In the future, improvements in sleep detection accuracy and more sophisticated hardware and software providing reliable feedback output function, may make it possible for the user to be warned of a possible sleep disorder, enabling earlier medical intervention. We also suggest testing future devices in sleep disordered and other special populations.

CITATION

Poyares D, Hirotsu C, Tufik S. Fitness tracker to assess sleep: beyond the market. *SLEEP* 2015;38(9):1351–1352.

DISCLOSURE STATEMENT

Dr. Poyares has participated in speaking engagements for Sanofi-Aventis and EMS and has consulted for Vitalaire Air Liquide. The other authors have indicated no financial conflicts of interest.

REFERENCES

1. de Zambotti M, Baker FC, Colrain IM. Validation of sleep-tracking technology compared with polysomnography in adolescents. Sleep 2015;38:1461–8.
2. Lee J, Finkelstein J. Activity trackers: a critical review. Stud Health Technol Inform 2014;205:558–62.
3. Finkelstein J, Knight A, Marinopoulos S, et al. Enabling Patient-Centered Care Through Health Information Technology. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012 Jun. (Evidence Reports/Technology Assessments, No. 206.) Available from: http://www.ncbi.nlm.nih.gov/books/NBK99854/.
4. de Souza L, Benedito-Silva AA, Pires ML, Poyares D, Tufik S, Calil HM. Further validation of actigraphy for sleep studies. Sleep 2003;26:81–5.
5. Marino M, Li Y, Rueschman MN, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. Sleep 2013;36:1747–55.
6. Montgomery-Downs HE, Insana SP, Bond JA. Movement toward a novel activity monitoring device. Sleep Breath 2012;16:913–7.
7. 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;11:709–15.
8. 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;9:1–5.
9. Kim J. Analysis of health consumers' behavior using self-tracker for activity, sleep, and diet. Telemed J E Health 2014;20:552–8.
10. Kim J. A qualitative analysis of user experiences with a self-tracker for activity, sleep, and diet. Interact J Med Res 2014;3:e8.