Consumer Sleep Apps: When it Comes to the Big Picture, it's All About the Frame

Commentary on Bhat 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.

Matt T. Bianchi, MD, PhD

Neurology Department, Sleep Division, Massachusetts General Hospital, Boston MA

While the consumer space is overflowing with new technologies for monitoring sleep, the flow of corresponding validation data is more like a trickle. The optimistic response to consumer sleep tracking market growth is to celebrate the public's increased awareness and interest in this critical aspect of our health and well-being. On the other hand, given the blurred boundary between sleep as "wellness" and sleep as a medical field, one cannot ignore important questions regarding the (mis)alignment between marketing claims, consumer expectations, experimental validation, and clinical utility. Reflexively posing the almost rhetorical question of whether consumer devices are as accurate as polysomnography (PSG) conjures images of babies and bathwater. Reframing the question may be the most useful path to answering whether, and how, these products fit into sleep medicine and sleep research.

Bhat et al. take an important step toward bringing the conversation about consumer sleep apps into the realm of independent validation studies.[1] They undertook a simple protocol to ask how well the output metrics of a popular smart phone app (Azumio) matched the sleep-wake staging of concurrent PSG. Anyone familiar with movement-based staging algorithms used in actigraphy monitoring will find no surprise that the phone-near-pillow app cannot distinguish sleep stages. In fact, so extensive is the published experience with actigraphy algorithms, that any other outcome would have been viewed with great Bayesian skepticism. In contrast, and perhaps surprisingly, the sleep versus wake discrimination was reasonable and in fact quite similar to that reported for wrist actigraphy: ~90% sensitive and ~50% specific for sleep.[2,3] Like wrist actigraphy, the app overestimated sleep, probably because quiet wakefulness contains little movement. Given the lack of distinction between sleep sub-stages, it is not surprising that the smart-alarm function was ineffective in this report.

One key validation challenge was the lack of access to the raw data, forcing the authors to manually extract the app staging data in epochs of much larger duration than used clinically. Raw data access is also crucial because new algorithms are continually being developed that can enhance information extraction, even from data as apparently simple as limb movement. For example, recent work illustrates that advanced transition probability analysis of actigraphy offers important phenotypic insights into sleep disruption.[4]

This app validation work adds to a discussion still in its infancy about the role of consumer monitors in sleep research and practice.[5] Although appropriately cautious given the poor correlation with sleep sub-stages, Bhat et al. provide a broader perspective to consider possible frameworks in which consumer monitors could be used. For the Azumio, the niche may be essentially that of traditional wrist actigraphy, where the field has already accepted the limitations while still cogently delineating a framework for appropriateness of tracking gross sleep-wake patterns.[6] Movement-based tracking can provide adjunctive data to manual diaries for long-term monitoring, or may supplant diaries in those who cannot (or will not) adhere to self-reporting. It can also complement diary responses as outcomes for intervention studies, and can even serve as a form of biofeedback, as reported previously for patients with misperception insomnia.[7]

The issue of external validity is especially important in the field of sleep medicine, where physiology and symptoms often dissociate, occult primary sleep problems are not uncommon, and physiology depends on state and trait alike. For example, the formerly available Zeo monitor published reasonable accuracy of sleep stage discrimination in healthy adults.[8] However, performance was negatively impacted by bruxism[9] and by neuroactive medications (unpublished results), which altered the forehead-based signal properties and thus the algorithm output accuracy.

The related issue of ecological validity is perhaps even more challenging for evaluating smart-phone apps as an inexpensive alternative to formal actigraphy. Larger samples with longitudinal monitoring, using actigraphy as the gold standard, are sorely needed to explore the parameter space of "real world" tracking. For example, the position of the phone, the movements of a bed partner, the type of mattress, and other factors may influence app sensitivity and specificity (which, despite the commonly held assumption, are not inherent properties of a diagnostic tool[10]).

Important conceptual challenges for this and many other consumer sleep products relate to the marketing content. The Azumio website claims that the app has a smart-alarm feature to wake you "at the perfect moment," and while it does not claim to be validated for this or for sleep staging, it does state that they "work closely with scientists and academic

institutions." This language may seem vague to the scientist, but may nevertheless sound authoritative to the uninitiated. Additionally, even the common practice of using terminology of "light" versus "deep" sleep, which may carry vernacular implications of bad versus good sleep, risks creating the false impression of health implication of the staging output.

Beyond the need to create a culture that views consumer product validation as a goal rather than a burden, creative solutions are needed to set consumer expectations and to engage companies in the importance of external and ecological validity. Consumers currently encounter far more "validation" from product reviews than from experimental findings. Aligning user expectations, marketing content, and validation data may also avoid the potential for "backlash" against individual devices or even consumer sleep monitors in general; at least one company is currently under fire for its sleep tracking claims.[11] While the disclaimers of "not a medical device" may seem innocent enough, the challenge remains that a consumer may not know whether their sleep concerns ought to reside in the realm of wellness or in the realm of medicine, and for some the app data might be their only guide. Few of the consumer sleep products have pursued FDA pathways, which may create further challenges for providers who, despite their interest in potential clinical uses, remain on the sidelines pending the legitimacy provided by FDA clearance.

As more groups pursue validation studies, and more companies recognize ways that they can contribute, even if simply through access to raw data, we can more appropriately frame questions of consumer sleep monitor utility. Stakeholders should be prepared, however, for the likely but unwieldy outcome of broader validation pursuits: that certain devices might be useful, in certain circumstances, for certain kinds of people, to answer certain kinds of questions.

CITATION

Bianchi MT. Consumer sleep apps: when it comes to the big picture, it's all about the frame. *J Clin Sleep Med* 2015;11(7):695–696.

REFERENCES

1. 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.
2. Martin JL, Hakim AD. Wrist actigraphy. Chest 2011;139:1514–27.
3. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003;26:342–92.
4. Lim AS, Yu L, Costa MD, et al. Quantification of the fragmentation of rest-activity patterns in elderly individuals using a state transition analysis. Sleep 2011;34:1569–81.
5. Behar J, Roebuck A, Domingos JS, Gederi E, Clifford GD. A review of current sleep screening applications for smartphones. Physiol Meas 2013;34:R29–46.
6. Littner M, Kushida CA, Anderson WM, et al. Practice parameters for the role of actigraphy in the study of sleep and circadian rhythms: an update for 2002. Sleep 2003;26:337–41.
7. Tang NK, Harvey AG. Altering misperception of sleep in insomnia: behavioral experiment versus verbal feedback. J Consult Clin Psychol 2006;74:767–76.
8. Shambroom JR, Fabregas SE, Johnstone J. Validation of an automated wireless system to monitor sleep in healthy adults. J Sleep Res 2012;21:221–30.
9. Kudesia RS, Bianchi MT. Decreased nocturnal awakenings in young adults performing bikram yoga: a low-constraint home sleep monitoring study. ISRN Neurology 2012;2012:153745.
10. Bianchi MT, Alexander BM. Evidence based diagnosis: does the language reflect the theory? BMJ 2006;333:442–5.
11. Comstock J. 2015; http://mobihealthnews.com/43499/class-action-lawsuit-alleges-fitbit-misled-buyers-with-inaccurate-sleep-tracking/. Accessed 5/17/2015.

SUBMISSION & CORRESPONDENCE INFORMATION

DISCLOSURE STATEMENT

Dr. Bianchi has consulted for Foramis, GrandRounds, Expert Testimony, and MC10; has received research support from MC10 and travel funds from Servier; and owns intellectual property rights associated with Rest Devices.