JCSM | Journal *of* Clinical Sleep Medicine

SPECIAL ARTICLES

Evaluating consumer and clinical sleep technologies: an American Academy of Sleep Medicine update

Sharon Schutte-Rodin, MD[1]; Maryann C. Deak, MD[2]; Seema Khosla, MD[3]; Cathy A. Goldstein, MD[4]; Michael Yurcheshen, MD[5]; Ambrose Chiang, MD[6]; Dominic Gault, MD[7]; Joseph Kern, MD[8]; Daniel O'Hearn, MD[9]; Scott Ryals, MD[10]; Nitun Verma, MD, MBA[11]; Douglas B. Kirsch, MD[12]; Kelly Baron, PhD[13]; Steven Holfinger, MD[14]; Jennifer Miller, APRN, PhD[15]; Ruchir Patel, MBBCh, MD[16]; Sumit Bhargava, MD[17]; Kannan Ramar, MD[18]

[1]University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; [2]eviCore Healthcare, Boston, Massachusetts; [3]North Dakota Center for Sleep, Fargo, North Dakota; [4]University of Michigan Sleep Disorders Center; [5]University of Rochester Medicine Sleep Center, Rochester, New York; [6]Louis Stokes Cleveland VA Medical Center, Case Western Reserve University, Cleveland, Ohio; [7]Greenville Health System, University of South Carolina, Greenville, South Carolina; [8]New Mexico VA Health Care System, Albuquerque, New Mexico; [9]Department of Medicine, University of Washington, Seattle, Washington; [10]University of Florida Health Sleep Center, Gainesville, Florida; [11]AC Wellness, Fremont, California; [12]Carolinas Healthcare Medical Group Sleep Services, Charlotte, North Carolina; [13]Univeristy of Utah Sleep-Wake Center, Salt Lake City, Utah; [14]Ohio Sate University Medical Center, Columbus, Ohio; [15]University of Nebraska Medical Center, Lincoln, Nebraska; [16]The Insomnia and Sleep Institute of Arizona, Scottsdale, Arizona; [17]Lucille Packard Children's Hospital at Stanford, Palo Alto, California; [18]Mayo Clinic, Rochester, Minnesota

INTRODUCTION

The previous American Academy of Sleep Medicine (AASM) Consumer and Clinical Technology Committee and the subsequent Technology Innovation Committee have assessed a wide range of consumer and clinical sleep technologies. These assessments are available for AASM members on the AASM website (#SleepTechnology, found in the clinical resources section).[1] Sleep device/application (app) assessments include product claimed capabilities, narrative summaries, Food and Drug Administration (FDA) status, sensors, mechanisms, data outputs, raw data if available, application programming interfaces (APIs) if accessible, similar devices/apps, and if there are peer-reviewed validation studies or clinical trials. Assessments are intended to assist members in general product understanding and do not represent product endorsement. An updated list of commonly used sleep (device/app) technology terms, in order from simple to more complex terms, is provided in **Table 1**.

Over the past few years, we have seen the evolution of some consumer sleep technologies used for self-tracking and/or self-help, to devices/apps with potentially meaningful clinical diagnostic, therapeutic, and/or long-term data tracking uses. See **Figure 1** and the following descriptions of types of sleep device/app technologies.

Consumer grade technologies generally promote sleep self-awareness and/or may provide suggestions for improving sleep. Definitions of metrics like "sleep quality" or "sleep scores" may vary between devices/apps. Consumer devices/apps do not require a prescription. They may or may not be registered for FDA wellness and sports use.[2] Popular sleep/wake tracking smartwatches are an example of consumer long-term sleep-wake self-tracking.

Clinical grade technologies require a prescription, are often FDA cleared or approved, and typically have some peer-reviewed validation studies. Continuous positive airway pressure (CPAP) data tracking by providers is an example of long-term remote patient monitoring (RPM), which offers enhanced care by providing clinical data monitoring between visits. By providing interim care, some RPM may be reimbursable.[3–5] Clinical grade technologies often utilize product-specific, proprietary artificial intelligence (AI), machine learning (ML), or deep learning (DL) algorithms.

Hybrid and/or transitional technologies represent an array of technologies. A **hybrid** technology may represent a technology for which one of its sensors is FDA cleared or approved for a specific metric, but whose overall claimed data output has not been validated or FDA cleared or approved. A **transitional** technology may be one that is in the process of applying for FDA clearance or approval using preliminary (often internal, unpublished) studies. Hybrid and transitional technologies frequently utilize product-specific, proprietary AI/ML/DL algorithms.

Previous statements developed by the AASM Consumer and Clinical Technology Committee recognized potential benefits, limitations, and risks of sleep-related technology disruptions.[6,7] In the context of the fast pace of sleep technology development, sleep medicine professionals look forward to technological advancements while considering these developments thoughtfully and scientifically before embracing them. This noted, the previous rapid pace of sleep technology development has been propelled even faster by the unexpected 2020 expansion of consumer and provider interest in self-monitoring physiological

Table 1—Commonly used sleep device/app technology terms.

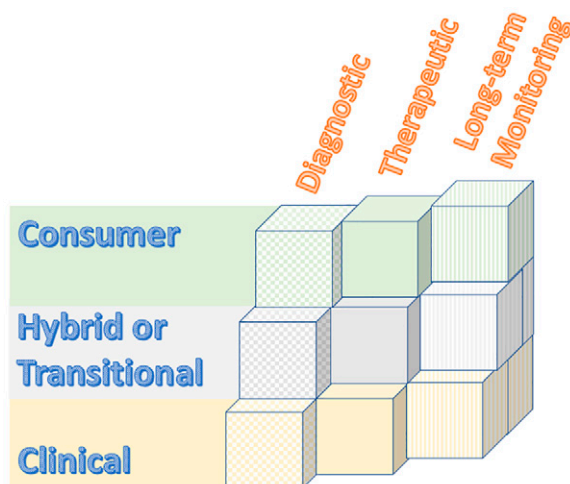|	
Wearable	Devices that are worn to provide physical data or feedback
Nearable	Nearby contactless devices that provide physiologic or environmental data or feedback
Sensor	A device that measures a physical input and converts it into understandable data
Photoplethysmography (PPG)	PPG sensors use a light source and a photodetector to measure blood flow changes, which provide signals that may use AI/ML/DL algorithms to deliver data outputs such as sleep stages
Sleep score or quality	Often a product specific computation of "sleep quality" derived from questionnaires and/or sensor data
Sleep stages	Device/app reporting of sleep stages such as "light sleep" or "deep sleep" that may vary in type of data acquisition, derivations, and definitions between devices/apps; staging may be derived from proprietary AI/ML/DL algorithms such as using PPG heart rate variability (HRV) rather than standard polysomnographic EEG scoring rules
Patient generated health data (PGHD)	Health care related data that are generated by patients and collected for the purpose to address a health concern or issue
Mhealth (mobile health)	The use of mobile phones or other wireless technologies to monitor and exchange health information
Software as a medical device (SaMD)	Software intended for medical uses that does so without being part of a hardware medical device
Mobile medical app (MMA)	A mobile app whose functionality meets the definition of a medical device
Clinical decision support (CDS) software	SaMD software risk categorization established by the International Medical Device Regulators Forum to determine if a software treats, diagnoses, or drives or informs clinical management
Artificial intelligence/machine learning (AI/ML) as an SaMD	SaMD that may have "locked" AI/ML algorithms or "adaptive learning" algorithms that may be assessed using an FDA Precert total Lifecycle product approach
Remote data monitoring	Monitoring of data remotely
Remote patient monitoring	A subset of remote data monitoring that is used clinically
Application programming interface (API)	A software interface that allows two or more applications to exchange information such as with an electronic health record
Algorithm	A sequence of statistical processing steps to solve a problem or compete a task
Artificial intelligence (AI)	The broad use of computer algorithms to simulate human tasks and thinking
Machine learning (ML)	A subset of AI that uses data training sets to make predictions and decisions without explicit programming
Deep learning (DL)	A subset of ML that enhances a deeper dive into smaller patterns of artificial neural networks

data, telemedicine, remote testing, remote data monitoring, and novel device sensor and AI/ML/DL integrations catalyzed by the COVID-19 pandemic.[8–11] Sleep technology sensors, such as pulse oximeters or EEG sensors, have been added to previous

Figure 1—Sleep device/app types.

consumer grade rings, watches, eye or head bands, and other wearables or nearables. Proprietary AI/ML/DL algorithms, such as those using heart rate variability (HRV), have assisted in the progression from consumer to transitional and clinical grade sleep technologies.

Thus, the need for clarification in how to evaluate rapidly evolving and diverse consumer and clinical sleep technology types and uses has become even more timely to sleep providers.[12] Specifically, confirming validation of product marketing claimed capabilities can be a challenge for busy clinicians who may be seeking peer-reviewed, randomized controlled studies for each device/app. However, such traditional validation studies may require too long of a time frame to complete and publish for real-time assessments of device/app performance or accuracy. For rapidly emerging technologies, novel validation processes may help lead to faster integration into clinical care applications. Examples may include outcomes-based or AI/ML/DL assisted validation. With these challenges in mind, we propose the following items for clinicians to consider when evaluating the wide range of consumer, hybrid or transitional, and clinical sleep-related technologies:

- Awareness of FDA terms
- Defining sleep term definitions across devices/apps
- Defining populations

- Data integrity
- Applications of new sensors, new sensor applications, or other novel technologies
- Awareness of proprietary AI/ML/DL algorithms
- Defining validation methods for claimed capabilities

AWARENESS OF FDA TERMS

Evaluation of sleep devices/apps is enhanced by an understanding of FDA terms. Unless the product has a specific exemption, FDA classification is generally based on the device/app safety risk, the intended use, and the indication(s) for use.[13,14] **Premarket notification, or 510(k) FDA clearance** for marketing, allows the FDA to determine if the product is equivalent to a "predicate" device/app already placed in Class I (low risk), Class II (moderate or higher risk than Class I), or Class III (high risk) type category.[15] 510(k) clearance is often required for Class II devices/apps and does not require clinical trials. The device/app requires: (1) Having the same intended use AND technological characteristics as the predicate device OR (2) Having the same intended use AND different safe and effective technology. The FDA 510(k) gives marketing clearance (but not "approval") to such Class I or II devices. **Premarket approval (PMA)**, typically used for Class III devices and more in-depth than 510(k), requires that the device is safe and effective. The PMA process generally requires human clinical trials supported with lab testing.[16] **FDA approval** requires successful PMA or a specific exemption before a device can be marketed. **FDA Granted** is a new term by which the device/app uses the **De Novo pathway** before it can be marketed. The De Novo pathway is offered for Class I or II devices with low to moderate safety risk when there is no similar predicate device.[17]

General Wellness devices/apps do not require FDA 510(k) clearance or PMA approval, and they may or may not be FDA registered or require enforcement discretion.[18] These Wellness and Sports devices/apps: (1) Have low safety risks to the user or others, (2) Are intended for wellness purposes, (3) Do not directly purport to diagnose or treat a disease or medical condition, and (4) Do not require FDA clearance or approval. Of special note, some popular health/sleep watches may be wellness type or may have FDA 510(k) clearance as a Class I or II medical device. The FDA also has provided guidance on **mobile medical apps** (MMA).[19] The FDA does not have a policy for the storage or platforms for these apps. However, some mobile apps may fall under the FDA **software as a medical device (SaMD)**. The FDA provides guidance for SaMD which are defined as software used for medical purposes that do so without being part of a hardware medical device.[20] Further, through the FDA Digital Health Center of Excellence, the FDA is providing guidance and advanced digital pathways for other SaMD type apps such as the International Medical Device Regulators Forum risk categories for **clinical decision support (CDS)** software, the FDA patient-generated health data throughout the **total product life cycle (TPLC)** device/app approach, and the **AI/ML as SaMD** approach.[21] Like the other FDA designations, these device/app pathways are guided by safety risk, intended use, and indications for use. **Devices@FDA** provides one place where you can find official information about FDA 510(k) cleared and PMA approved medical devices/apps. This includes summaries of currently marketed medical devices/apps.

DEFINING SLEEP DATA TERMS ACROSS DEVICES/APPS

The use of multiple or nonspecific definitions for sleep-related terms may be encountered when reviewing sleep devices/apps. For example, sleep quality, sleep scores, light and deep sleep, and other terms may have variable definitions across devices/apps. Ideally, terminology should be specifically defined and reproducible across comparable consumer or clinical output metrics. Some commonly used sleep device/app and technology terms are listed in **Table 1**.

In addition, some platforms claim to integrate multiple consumer and/or clinical data outputs.[22,23] While integrated platforms may allow easier data and report access, terms across devices/apps may carry different meanings, which may cause confusion particularly if using an integrated platform with multiple data sources and sleep term definitions. For example, definitions, derivations, and outputs of "light sleep" may vary across devices/apps.

Additionally, integration of consumer or clinical device/app data directly into the patient's electronic health record (EHR), a legal medical document, also raises questions about using consistent sleep terms, device/app data validation, who reviews and owns these vast newly available datasets, and accepted standards for uses of such data within the EHR. Health care disciplines have come to expect standardized definitions when products are used for research or clinical purposes.[24,25] More relevant to the sleep field, the Consumer Technology Association (CTA) and the National Sleep Foundation (NSF) have produced three white papers specific to defining device/app terminologies for researchers, clinicians, industry, and consumers. Developed jointly by CTA and the NSF, Definitions and Characteristics for Wearable Sleep Monitors (2016) and Methodology of Measurements for Features in Sleep Tracking Consumer Technology Devices and Applications (2017) provide a foundation for defining sleep terms for use across different sleep apps and devices.[26,27] These groups discuss standardizing sleep metrics in terms of events, processes, and patterns in Performance Criteria and Testing Protocols for Features in Sleep Tracking Consumer Technology Devices and Applications (2019).[28] This working group is an example of how clinicians, researchers, developers, and industry can benefit from collaboration.[29]

DEFINING POPULATIONS

Validated and/or FDA cleared/approved devices and apps may not necessarily be validated for all age ranges, or for populations with sleep disorders and/or medical comorbidities, or for patients taking certain medications or having implantable devices. Sleep "best practices," accepted standards of care, quality measures, and clinical guidelines include indications and contraindications for testing of specified populations.[30–39] For example, home sleep apnea testing (HSAT) is generally indicated for patients who have been prescreened for uncomplicated obstructive sleep apnea

(OSA) testing (such as patients without significant cardiopulmonary comorbidities). A sleep device/app should indicate if its claimed use is for a particular consumer or patient population, as well as cite any exclusionary populations. For technologies that will be used to diagnose sleep disorders, direct treatment decisions, or drive personal or population health, clear definitions of the claimed uses for specified populations are indicated. Population demographics of the accessible datasets should accompany these devices/platforms to inform generalizability. These minimum, general requirements are much like the FDA's digital health action and innovation plans.[40,41]

DATA INTEGRITY

Considerations for the integrity of patient generated health data (PGHD) include: who may access device/app data, privacy, ownership, storage, security, raw data review, API accessibility or appropriate platform integration (such as with an EHR or with other entities). Raw data review and interpretation may be meaningful features to some sleep providers, as in the case of reviewing the raw data of apnea testing devices. Who may view, own, or share the data, may be other important features to users. Approximately 45% of smartphone devices have health or fitness apps,[42] and the security of PGHD and personal health information has received recent attention.[43] Health care entities appear to be a favored target for hackers.[44] In a 2020 security report on global mHealth apps, a leading app security firm reported that 71% of health care and medical apps had vulnerability, and 91% had weak encryption.[45,46] APIs are reported to be highly vulnerable to hacking, including EHR access.[47] Of current relevance, the 21st Century Cures Act has directed that patients are able to access their EHR data as of April 2021, often using APIs.[48] Also relevant, the Office of the National Coordinator for Health Information Technology (ONC) supports using standardized APIs such as the Fast Healthcare Interoperability Resources (FHIR) and endorses eradicating information blocking.[49,50] The American Medical Association has provided guidelines about protecting PGHD.[51] Further, the bidirectional information flow from and to a medical device is a consideration. While there are no current reports of hacking sleep medical devices, the FDA has warned about the ability to hack medical devices such as pacemakers, insulin pumps, and other devices.[52–54]

APPLICATIONS OF NEW SENSORS, NEW SENSOR APPLICATIONS, OR OTHER NOVEL TECHNOLOGIES

Wearable sleep sensors may collect data using finger probes, nasal/oral/mask sources, rings, watches/wrist bands, torso bands, skin patches, eyewear, forehead or headbands or caps, smart garments, shoe inserts, leg bands, or other worn devices.[55,56] **Nearables** often are located on a bedside stand or under the mattress. Types of nearable sleep tracking include ballistocardiography vibration (for respiratory and heart rate, stroke volume), sound, light, temperature, humidity, and movement sensors.[57–60] Like wearables, nearables often report user snoring, sleep times, staging, and "quality."[61]

Sleep clinicians and researchers are familiar with triaxial accelerometers for actigraphy or polysomnographic (PSG) sensors such as EEG, EOG, ECG, EMG, nasal/oral airflow sensors, torso belts, pulse oximeter, microphone, and camera. Newer consumer and clinical sleep technologies may utilize combinations of these sensors, other sensors, or novel applications of sensor data that may increase the accuracy or performance of the output data such as sleep staging.[62,63] Examples of other sensors include skin temperature, radar/radiofrequency, sound, environmental sensors, ingestibles, and others.[63–65]

Further, using AI/ML/DL, some sensor data has been transformed and extended for other uses and applications. For example, like a pulse oximeter, a photoplethysmography (PPG) device uses a light source (through vascular tissue with pulsatile blood-volume flow) and an oppositely positioned photodetector to measure changes in light intensity. The light source wavelength(s) and its application type and location are often product specific. The PPG waveform consists of an AC current (pulsatile wave) superimposed on a DC current (steady, slow changes with respiration). Sensitive to motion and other artifact, raw PPG waveforms are amplified, filtered, and derived to provide outputs such as heart rate variability (HRV) or peripheral arterial tone (PAT).[66,67] Using datasets and these signals, AI/ML/DL algorithms are used to provide users with familiar data such as apnea-hypopnea index (AHI) or sleep stages.[68–76] Note that pulse oximeters may be affected by position on the hemoglobin-oxygen desaturation curve, poor pulsatile flow, hypoxia, motion, skin tone, and other settings.[77–79] Similarly, products utilizing PPG may have variable accuracy in different conditions and populations. Additionally, a product algorithm is specific for a sensor at a specified location, type of light source, and its type of application and data acquisition. For example, data derived from one fingertip sensor brand cannot be generalized to another device's fingertip or wrist application.

Many newer clinical, hybrid, and evolving transitional sleep technologies incorporate triaxial accelerometer and PPG sensors and utilize proprietary AI/ML/DL algorithms to report sleep-related data.[62] However, each device generally utilizes different sensor designs, data collection and analysis methods, and AI/ML/DL data interpretation. Thus, sensor collection, location, analysis, and outputs such as PPG may differ across devices. As such, assessing validation or performance of claimed uses may be proprietary, product specific, and not generalizable across other devices. Proprietary sensors and associated algorithmic outputs are device/app specific.

AWARENESS OF PROPRIETARY AI/ML/DL ALGORITHMS

Sensors allow the collection of vast amounts of physiologic and environmental data, and AI-based analytics are well suited to present digestible data outputs and summaries to patients, providers, and researchers. To display user-friendly data output, sleep device/app software may use AI/ML/DL to incorporate consumer or patient-entered data and data from one or more sensors.[80–82] Based on ongoing user entered and/or physiological data collection, some AI/ML/DL software may advise or make

patient-centered, focused recommendations to a consumer or to a patient. For example, CPAP or insomnia software may provide coaching based on collected data. Using learning datasets, AI/ML/DL algorithms learn from data to improve performance on specific tasks (for example, automated sleep staging). Such computerized algorithms are often proprietary, not easily summarized, and may be referred to as a "black box."[83] However, even while remaining proprietary, disclosure of certain aspects of algorithm development and testing can improve transparency and, therefore, increase confidence in the clinical use of such tools.[84,85]

When AI/ML/DL analysis is applied to a new technology to track sleep, information regarding both the training and testing dataset should be disclosed including the size, demographics, and clinical features of participants from which the data is derived. Examples of shared sleep datasets are available.[86]

As the first clinical application of AI/ML/DL in sleep medicine is likely to be the automated scoring of sleep and associated events, a certification program could provide a framework to guide sleep laboratories in the use of ML-based scoring software. In addition to disclosure of minimum characteristics of training and testing sets and reporting of performance of statistics on a novel, independent testing set, such a program will also require manufacturers to assist sleep laboratories with demonstration of local performance in their own facility.

DEFINING VALIDATION METHODS FOR THE DEVICE/APP CLAIMED CAPABILITIES

Providers and researchers commonly expect clinical (diagnostic or treatment) validation studies to be peer-reviewed, gold standard comparisons, randomized clinical trials, and/or outcomes-based or population health-based strategies. Clinicians generally use the requirement of a prescription or some type of FDA verbiage to guide comfort in using a product. This noted, users may encounter difficulties when searching for support or validation of product marketing statements about claimed uses.[87] Additionally, sleep stage validation for consumer or clinical sleep devices against gold standards may be an indirect comparison of different sensors and interpretive derivations. For example, the former may use proprietary AI/ML/DL processed heart rate signals compared against polysomnographic sensor data using EEG, EOG, EMG.[25]

Determining the accuracy or performance of sleep device/app claimed uses can be challenging for a variety of reasons:

- A descriptive name for a sensor may vary in accuracy across specific devices. For example, the accuracy of pulse oximeters may vary across products.
- Sensor or device position and/or environment may affect data collection.
- Validation for the claimed uses may not specify inclusion/exclusion for the intended population use (such as for certain age ranges, healthy vs users with sleep or medical comorbidities, users with pacemakers or taking certain medications, other conditions).
- Cited articles may be exclusively performed and/or funded by the product company and/or inventor.
- Cited articles may refer to general sleep principles but lack specific validation for that particular product claim. For

example, the referencing of general light wavelength and circadian studies does not verify that a particular light-related device/app has authenticated the claimed use for that specific light-related device/app.

- Cited validation articles may be for one sensor, but not for the integration with other sensors or the associated proprietary AI/ML/DL algorithms for that product claim. Validation or FDA clearance of one sensor does not necessarily ensure fidelity of the entire technology, which requires sound methods of data collection for all incorporated sensors as well as data transmission and analysis methods. Some devices/apps utilizing proprietary "black box" algorithms also may incorporate patient-reported data and demographics with data from multiple sensors (such as photoplethysmography, accelerometer, electrocardiogram, pulse oximetry).
- Software and/or proprietary AI/ML/DL algorithms may be adjusted and/or updated, with or without specific notice to users.[25] Additionally, even when validated studies may be available, the testing may have been performed on an outdated software version or dataset.

For developers who design products intended for medical and research purposes, traditional "gold standard" comparison validation studies may not be applicable or may be difficult to achieve in a brisk marketplace. This noted, transparent validation efforts and guidelines for choosing sleep technologies clinically or for research are needed.[83] For example, validation may not have to be limited to traditional "gold standard" comparisons, but rather could include clinical certification based on comparative outcome studies, AI/ML/DL validation using shared datasets, and/or other novel validation approaches. Menghini et al recently described a standardized framework for testing the performance of sleep trackers.[88]

CONCLUSIONS

Like all fields of medicine, sleep medicine is in an era of rapid technologic disruption that has been further accelerated by recent care delivery changes, telehealth expansion, and remote data monitoring accompanying the COVID-19 pandemic. Also catalyzed by the pandemic, consumers have sought devices/apps that allow sleep self-help measures and data tracking. The influx of new sensors and mobile applications, new modes of remote signal collection, and new methods of data management are generating massive information caches. Analytics with AI/ML/DL could realize the promise of personalized medicine, precision medicine, and population health.

There are a broad range of consumer, hybrid and transitional, and clinical sleep technologies, and providers look to deliver a consistent message about this range of sleep technologies to patients. Consumer sleep technologies encourage patients to think about their sleep and its impact on their health. Such discussions open conversations about the patients' sleep concerns, questions, or symptoms.

While PSG is considered the gold standard for defining sleep stage metrics and diagnosing many sleep disorders, PSG is

limited by its data collection outside of the home and for it providing only a snapshot in time. Home sleep apnea tests also provide snapshot data, do not typically include arousals, and may have variable accuracy of sleep metrics across devices. Actigraphy too has limits of its data acquisition being expensive and time-consuming as well as data collection generally for only 2 weeks. Reliable consumer sleep technology could provide popular, inexpensive, 24/7 sleep/wake data collection over long periods of time.[89] New pathways for diagnostic testing, clinical treatments, and/or chronic management could emerge from such long-term data collection and analysis.

Sleep providers are familiar with the improvement in CPAP compliance with remote data monitoring and patient engagement software.[90–92] Providers look forward to adding enhanced remote testing, treatment, and both consumer and clinical data monitoring sleep technology tools to provide real-time and improved between-visit care, personalized care, interactive data alerts, and novel care guidelines based on personalized or population health big data analytics. AI/ML/DL analytics of ongoing patient reported data and physiologic data over time offer new opportunities to monitor individual and group patient symptoms and physiology in the home continuously, indefinitely, and in real time. Changes from baseline data trends could prove invaluable in predicting individual and group disease onset and/or exacerbations.[9,10,93,94]

However, before using sleep devices/apps clinically, providers seek to gain comfort in understanding how to assess the accuracy, performance, and intended uses of the product marketing claims. As more devices/apps utilize proprietary AI/ML/DL algorithms, user confidence is further challenged. Wu et al studied the 130 medical AI devices cleared by the FDA between 2015 and 2020 and found the FDA AI process less vigorous than the FDA pharmaceutical process.[95] As an example, the authors used a chest X-ray detection of pneumothorax algorithm and found that it worked well for the original site cohort but was 10% less accurate for two different sites and less accurate for Black patients. They cite that many datasets are retrospective, from only one or a few sites, or do not include all representative populations in clinical settings. Interested clinicians can search the FDA web database for devices/apps that have obtained 510(k) clearance or premarket approval. Yet, seekers cannot easily do similar FDA web AI/ML/DL searches, and Benjamens et al have created an open access database of strictly AI/ML-based medical technologies that have been approved by the FDA.[96]

As described in this paper (and **Table 2**), a practical checklist for clinicians when evaluating sleep product claimed uses includes: awareness of FDA terms, familiarity with product sleep term definitions, use with particular populations, data integrity considerations, recognizing sensor types and applications, awareness of proprietary AI/ML/DL algorithms, and clarification of validation methods used for the product claimed uses. At present, evaluating product claimed uses requires time-consuming verification and/or familiarity for each unique device/app. Until there are consistent performance standards for these elements across devices/apps, providers may continue to feel unsure about clinical use of the many and diverse sleep technologies. Recent guides have been proposed for developers to use to document the performance of product claimed uses, and consumers, clinicians, and researchers look forward to this consistency and

transparency. To provide user confidence, increased communications and collaboration between industry, government, insurers, clinicians, researchers, and consumers could help to create a standard framework for reporting of product performance.

Table 2—A guide to evaluating sleep technologies.

- Awareness of FDA terms
- Defining sleep term definitions across devices/apps
- Defining populations
- Data integrity
- Applications of new sensors, new sensor applications, or other novel technologies
- Awareness of proprietary AI/ML/DL algorithms
- Defining validation methods for claimed capabilities

REFERENCES

1. American Academy of Sleep Medicine. #SleepTechnology. https://aasm.org/consumer-clinical-sleep-technology/. Accessed March 12, 2021.

2. U.S. Food & Drug Administration. General Wellness: Policy for Low Risk Devices. https://www.fda.gov/media/90652/download. Accessed March 12, 2021.

3. American College of Physicians. Telehealth Guidance and Resources. https://www.acponline.org/practice-resources/business-resources/telehealth/remote-patient-monitoring. Accessed March 12, 2021.

4. Federal Register. Centers for Medicare and Medicaid Services. https://www.federalregister.gov/agencies/centers-for-medicare-medicaid-services. Accessed March 12, 2021.

5. Wicklund E. CMS Clarifies 2021 PFS Reimbursements for Remote Patient Monitoring. https://mhealthintelligence.com/news/cms-clarifies-2021-pfs-reimbursements-for-remote-patient-monitoring. Published January 20, 2021. Accessed March 12, 2021.

6. Khosla S, Deak MC, Gault D, et al. Consumer sleep technology: an American Academy of Sleep Medicine position statement. *J Clin Sleep Med.* 2018;14(5):877–880.

7. Khosla S, Deak MC, Gault D, et al. Consumer sleep technologies: how to balance the promises of new technology with evidence-based medicine and clinical guidelines. *J Clin Sleep Med.* 2019;15(1):163–165.

8. Radin JM, Wineinger NE, Topol EJ, Steinhubl SR. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. *Lancet Digit Health.* 2020;2(2):e85–e93.

9. Zhu T, Watkinson P, Clifton DA. Smartwatch data help detect COVID-19. *Nat Biomed Eng.* 2020;4(12):1125–1127.

10. Quer G, Radin JM, Gadaleta M, et al. Wearable sensor data and self-reported symptoms for COVID-19 detection. *Nat Med.* 2021;27(1):73–77.

11. Chiquoine J. Staying Healthy at Home: How COVID-19 Is Changing the Wellness Industry. https://www.uschamber.com/co/good-company/launch-pad/pandemic-is-changing-wellness-industry. Accessed March 12, 2021.

12. Watson NF, Lawlor C, Raymann RJEM. Will consumer sleep technologies change the way we practice sleep medicine? *J Clin Sleep Med.* 2019;15(1):159–161.

13. U.S. Food & Drug Administration. Classify Your Medical Device. https://www.fda.gov/medical-devices/overview-device-regulation/classify-your-medical-device. Accessed March 12, 2021.

14. Piermatteo K. How Is My Medical Device Classified? https://fda.yorkcast.com/webcast/Play/17792840509f49f0875806b6e9a1be471d. Accessed March 12, 2021.

15. U.S. Food & Drug Administration. Premarket Notification 510(k). https://www.fda.gov/medical-devices/premarket-submissions/premarket-notification-510k. Accessed April 27, 2021.

16. U.S. Food & Drug Administration. Premarket Approval. https://www.fda.gov/medical-devices/premarket-submissions/premarket-approval-pma. Accessed April 27, 2021.

17. U.S. Food & Drug Administration. De Novo Classification Request. https://www.fda.gov/medical-devices/premarket-submissions/de-novo-classification-request. Accessed April 27, 2021

18. U.S. Food & Drug Administration. General Wellness: Policy for Low Risk Devices. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-wellness-policy-low-risk-devices. Accessed March 12, 2021.

19. U.S. Food & Drug Administration. Device Software Functions Including Mobile Medical Applications. https://www.fda.gov/medical-devices/digital-health-center-excellence/device-software-functions-including-mobile-medical-applications. Accessed March 12, 2021.

20. U.S. Food & Drug Administration. Software as a Medical Device (SaMD). https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd. Accessed March 12, 2021.

21. U.S. Food & Drug Administration. Digital Health Center of Excellence. https://www.fda.gov/medical-devices/digital-health-center-excellence. Accessed April 27, 2021.

22. Valdic. We Power Healthcare with Data. https://validic.com/. Accessed April 27, 2021.

23. Somnoware. One Platform. https://www.somnoware.com/. Accessed April 27, 2021.

24. Tison GH, Sanchez JM, Ballinger B, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. *JAMA Cardiol.* 2018;3(5):409–416.

25. de Zambotti M, Cellini N, Goldstone A, Colrain IM, Baker FC. Wearable sleep technology in clinical and research settings. *Med Sci Sports Exerc.* 2019;51(7):1538–1557.

26. Consumer Technology Association. Definitions and Characteristics for Wearable Sleep Monitors (ANSI/CTA/NSF-2052.1). https://shop.cta.tech/products/definitions-and-characteristics-for-wearable-sleep-monitors. Published September 2016. Accessed July 26, 2021.

27. Consumer Technology Association. Methodology of Measurements for Features in Sleep Tracking Consumer Technology Devices and Applications (ANSI/CTA/NSF-2052.2). https://shop.cta.tech/products/methodology-of-measurements-for-features-in-sleep-tracking-consumer-technology-devices-and-applications. Published September 2017. Accessed July 26, 2021.

28. Consumer Technology Association. Performance Criteria and Testing Protocols for Features in Sleep Tracking Consumer Technology Devices and Applications (ANSI/CTA/NSF-2052.3). https://shop.cta.tech/products/performance-criteria-and-testing-protocols-for-features-in-sleep-tracking-consumer-technology-devices-and-applications. Published April 2019. Accessed July 26, 2021.

29. Slotwiner DJ, Tarakji KG, Al-Khatib SM, et al. Transparent sharing of digital health data: a call to action. *Heart Rhythm.* 2019;16(9):e95–e106.

30. Kapur VK, Auckley DH, Chowdhuri S, et al. Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American Academy of Sleep Medicine clinical practice guideline. *J Clin Sleep Med.* 2017;13(3):479–504.

31. Rosen IM, Kirsch DB, Carden KA, et al. Clinical use of a home sleep apnea test: an updated American Academy of Sleep Medicine position statement. *J Clin Sleep Med.* 2018;14(12):2075–2077.

32. Aurora RN, Collop NA, Jacobowitz O, Thomas SM, Quan SF, Aronsky AJ. Quality measures for the care of adult patients with obstructive sleep apnea. *J Clin Sleep Med.* 2015;11(3):357–383.

33. Aurora RN, Quan SF. Quality measure for screening for adult obstructive sleep apnea by primary care physicians. *J Clin Sleep Med.* 2016;12(8):1185–1187.

34. Morgenthaler TI, Aronsky AJ, Carden KA, Chervin RD, Thomas SM, Watson NF. Measurement of quality to improve care in sleep medicine. *J Clin Sleep Med.* 2015;11(3):279–291.

35. Qaseem A, Holty JE, Owens DK, Dallas P, Starkey M, Shekelle P; Clinical Guidelines Committee of the American College of Physicians. Management of obstructive sleep apnea in adults: A clinical practice guideline from the American College of Physicians. *Ann Intern Med.* 2013;159(7):471–483.

36. Fleetham J, Ayas N, Bradley D, et al. Canadian Thoracic Society 2011 guideline update: diagnosis and treatment of sleep disordered breathing. *Can Respir J.* 2011;18(1):25–47.

37. Epstein LJ, Kristo D, Strollo PJ Jr, et al. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med.* 2009;5(3):263–276.

38. Collop NA, Anderson WM, Boehlecke B, et al. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. *J Clin Sleep Med.* 2007;3(7):737–747.

39. Kushida CA, Littner MR, Morgenthaler T, et al. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep.* 2005;28(4):499–523.

40. U.S. Food & Drug Administration. Global Approach to Software as a Medical Device. https://www.fda.gov/medical-devices/software-medical-device-samd/global-approach-software-medical-device. Accessed April 27, 2021.

41. U.S. Food & Drug Administration. Consumers (Medical Devices). https://www.fda.gov/medical-devices/resources-you-medical-devices/consumers-medical-devices. Accessed March 12, 2021.

42. Kunz T, Lange B, Selzer A. [Digital public health: data protection and data security]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2020;63(2):206–214.

43. Davis J. Top Health IT Security Challenges? Medical Devices, Cloud Security. https://healthitsecurity.com/news/top-health-it-security-challenges-medical-devices-cloud-security. Published January 14, 2021. Accessed April 27, 2021.

44. Davis J. Healthcare Accounts for 79% of All Reported Breaches, Attacks Rise 45%. https://healthitsecurity.com/news/healthcare-accounts-for-79-of-all-reported-breaches-attacks-rise-45. Published January 5, 2021. Accessed March 20, 2021.

45. Ghosh S. 71% of Healthcare Medical Apps Have a Serious Vulnerability; 91% Fail Crypto Tests. https://aithority.com/ait-featured-posts/71-of-healthcare-medical-apps-have-a-serious-vulnerability-91-fail-crypto-tests. Published September 29, 2020. Accessed March 20, 2021.

46. Security Magazine website. Cryptographic vulnerabilities, data leakage and other security breaches in healthcare apps. https://www.securitymagazine.com/articles/93524-cryptographic-vulnerabilities-data-leakage-and-other-security-breaches-in-healthcare-apps. Accessed March 20, 2021.

47. Horowitz BT. Mobile health apps leak sensitive data through APIs, report finds. https://www.fiercehealthcare.com/tech/mobile-health-apps-leak-sensitive-data-through-apis-report-finds. Published February 24, 2021. Accessed March 20, 2021.

48. U.S. Food & Drug Administration. 21st Century Cures Act. https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act. Accessed April 27, 2021.

49. Jason C. Patients Need Further Knowledge on APIs, Mobile Apps. https://ehrintelligence.com/news/patients-need-further-knowledge-on-apis-mobile-apps. Published November 30, 2020. Accessed April 27, 2021.

50. Jason C. API Adoption Can Accelerate Interoperability, Patient Data Exchange. https://ehrintelligence.com/news/api-adoption-can-accelerate-interoperability-patient-data-exchange. Published January 27, 2021. Accessed March 21, 2021.

51. Henry TA. How to keep patient information secure in mHealth apps. https://www.ama-assn.org/practice-management/digital/how-keep-patient-information-secure-mhealth-apps. Published January 13, 2020. Accessed March 21, 2021.

52. Jaret P. Exposing vulnerabilities: How hackers could target your medical devices. https://www.aamc.org/news-insights/exposing-vulnerabilities-how-hackers-could-target-your-medical-devices. Published November 12, 2018. Accessed March 21, 2021.

53. U.S. Food & Drug Administration. Cybersecurity. https://www.fda.gov/medical-devices/digital-health-center-excellence/cybersecurity. Accessed March 21, 2021.

54. U.S. Food & Drug Administration. Medical Device Cybersecurity: What You Need to Know. https://www.fda.gov/consumers/consumer-updates/medical-device-cybersecurity-what-you-need-know. Accessed March 21, 2021.

55. Labios L. New Skin Patch Brings Us Closer to Wearable, All-In-One Health Monitor. https://ucsdnews.ucsd.edu/pressrelease/new-skin-patch-brings-us-closer-to-wearable-all-in-one-health-monitor. Published February 15, 2021. Accessed March 23, 2021.

56. Shi H, Zhao H, Liu Y, Gao W, Dou S-C. Systematic Analysis of a Military Wearable Device Based on a Multi-Level Fusion Framework: Research Directions. *Sensors (Basel).* 2019;19(12):2651.

57. Albukhari A, Lima F, Mescheder U. Bed-embedded heart and respiration rates detection by longitudinal ballistocardiography and pattern recognition. *Sensors (Basel).* 2019;19(6):1451.

58. MuRata. Products and Sensors. https://www.murata.com/en-us/products/ Accessed March 23, 2021.

59. Bianchi MT. Sleep devices: wearables and nearables, informational and interventional, consumer and clinical. *Metabolism.* 2018;84:99–108.

60. Tuominen J, Peltola K, Saaresranta T, Valli K. Sleep parameter assessment accuracy of a consumer home sleep monitoring ballistocardiograph beddit sleep tracker: a validation study. *J Clin Sleep Med.* 2019;15(3):483–487.

61. Scott H. Sleep and circadian wearable technologies: considerations toward device validation and application. *Sleep.* 2020;43(12):zsaa163.

62. de Zambotti M, Cellini N, Menghini L, Sarlo M, Baker FC. Sensors capabilities, performance, and use of consumer sleep technology. *Sleep Med Clin.* 2020;15(1):1–30.

63. Chinoy ED, Cuellar JA, Huwa KE, et al. Performance of seven consumer sleep-tracking devices compared with polysomnography. *Sleep.* 2021;44(5): zsaa291.

64. Heglum HSA, Kallestad H, Vethe D, Langsrud K, Sand R, Engstrøm M. Distinguishing sleep from wake with a radar sensor A contact-free real-time sleep monitor. *Sleep.* 2021;zsab060.

65. Dafna E, Tarasiuk A, Zigel Y. Sleep staging using nocturnal sound analysis. *Sci Rep.* 2018;8(1):13474.

66. Tamura T. Current progress of photoplethysmography and SPO_2 for health monitoring. *Biomed Eng Lett.* 2019;9(1):21–36.

67. Castaneda D, Esparza A, Ghamari M, Soltanpur C, Nazeran H. A review on wearable photoplethysmography sensors and their potential future applications in health care. *Int J Biosens Bioelectron.* 2018;4(4):195–202.

68. Pang KP, Gourin CG, Terris DJ. A comparison of polysomnography and the WatchPAT in the diagnosis of obstructive sleep apnea. *Otolaryngol Head Neck Surg.* 2007;137(4):665–668.

69. Pillar G, Berall M, Berry R, et al. Detecting central sleep apnea in adult patients using WatchPAT-a multicenter validation study. *Sleep Breath.* 2020;24(1):387–398.

70. Massie F, Mendes de Almeida D, Dreesen P, Thijs I, Vranken J, Klerkx S. An evaluation of the NightOwl home sleep apnea testing system. *J Clin Sleep Med.* 2018; 14(10):1791–1796.

71. Jen R, Orr JE, Li Y, et al. Accuracy of WatchPAT for the diagnosis of obstructive sleep apnea in patients with chronic obstructive pulmonary disease. *COPD.* 2020;17(1):34–39.

72. Radha M, Fonseca P, Moreau A, et al. Sleep stage classification from heart-rate variability using long short-term memory neural networks. *Sci Rep.* 2019;9(1):14149.

73. Papini GB, Fonseca P, Gilst MMV, Bergmans JW, Vullings R, Overeem S. Respiratory activity extracted from wrist-worn reflective photoplethysmography in a sleep-disordered population. *Physiol Meas.* 2020;41(6):065010.

74. Papini GB, Fonseca P, van Gilst MM, Bergmans JWM, Vullings R, Overeem S. Wearable monitoring of sleep-disordered breathing: estimation of the apnea-hypopnea index using wrist-worn reflective photoplethysmography. *Sci Rep.* 2020;10(1):13512.

75. Papini GB, Fonseca P, van Gilst MM, et al. Estimation of the apnea-hypopnea index in a heterogeneous sleep-disordered population using optimised cardiovascular features. *Sci Rep.* 2019;9(1):17448.

76. Fonseca P, van Gilst MM, Radha M, et al. Automatic sleep staging using heart rate variability, body movements, and recurrent neural networks in a sleep disordered population. *Sleep.* 2020;43(9):zsaa048.

77. Luks AM, Swenson ER. Pulse oximetry for monitoring patients with COVID-19 at home. Potential pitfalls and practical guidance. *Ann Am Thorac Soc.* 2020;17(9): 1040–1046.

78. Baek HJ, Shin J, Cho J. The effect of optical crosstalk on accuracy of reflectance-type pulse oximeter for mobile healthcare. *J Healthc Eng.* 2018;2018:3521738.

79. Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. *Anesthesiology.* 2005;102(4):715–719.

80. Korkalainen H, Aakko J, Duce B, et al. Deep learning enables sleep staging from photoplethysmogram for patients with suspected sleep apnea. *Sleep.* 2020;43(11): zsaa098.

81. Walch O, Huang Y, Forger D, Goldstein C. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. *Sleep.* 2019;42(12):zsz180.

82. Fonseca P, Weysen T, Goelema MS, et al. Validation of photoplethysmography-based sleep staging compared with polysomnography in healthy middle-aged adults. *Sleep.* 2017;40(7).

83. Depner CM, Cheng PC, Devine JK, et al. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *Sleep.* 2020;43(2): zsz254.

84. Goldstein CA, Berry RB, Kent DT, et al. Artificial intelligence in sleep medicine: an American Academy of Sleep Medicine position statement. *J Clin Sleep Med.* 2020; 16(4):605–607.

85. Hwang TJ, Kesselheim AS, Vokinger KN. Lifecycle regulation of artificial intelligence- and machine learning-based software devices in medicine. *JAMA.* 2019;322(23): 2285–2286.

86. National Sleep Research Resource. https://sleepdata.org/. Accessed March 23, 2021.

87. Ko PR, Kientz JA, Choe EK, Kay M, Landis CA, Watson NF. Consumer sleep technologies: a review of the landscape. *J Clin Sleep Med.* 2015;11(12): 1455–1461.

88. Menghini L, Cellini N, Goldstone A, Baker FC, de Zambotti M. A standardized framework for testing the performance of sleep-tracking technology: step-by-step guidelines and open-source code. *Sleep.* 2021;44(2):zsaa170.

89. Goldstein CA, Depner C. Miles to go before we sleep … a step toward transparent evaluation of consumer sleep tracking devices. *Sleep.* 2021;44(2):zsab020.

90. Kuna ST, Shuttleworth D, Chi L, et al. Web-based access to positive airway pressure usage with or without an initial financial incentive improves treatment use in patients with obstructive sleep apnea. *Sleep.* 2015;38(8):1229–1236.

91. Malhotra A, Crocker ME, Willes L, Kelly C, Lynch S, Benjafield AV. Patient engagement using new technology to improve adherence to positive airway pressure therapy: a retrospective analysis. *Chest.* 2018;153(4):843–850.

92. Shaughnessy GF, Morgenthaler TI. The effect of patient-facing applications on positive airway pressure therapy adherence: a systematic review. *J Clin Sleep Med.* 2019;15(5):769–777.

93. Mishra T, Wang M, Metwally AA, et al. Pre-symptomatic detection of COVID-19 from smartwatch data. *Nat Biomed Eng.* 2020;4(12):1208–1220.

94. U.S. Food & Drug Administration. Coronavirus (COVID-19) Update: FDA Authorizes First Machine Learning-Based Screening Device to Identify Certain Biomarkers That May Indicate COVID-19 Infection. https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-first-machine-learning-based-screening-device-identify. Accessed March 23, 2021.

95. Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nat Med.* 2021;27(4):582–584.

96. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med.* 2020; 3(1):118.

SUBMISSION & CORRESPONDENCE INFORMATION

DISCLOSURE STATEMENT