Chapter

Novel Digital Wearable Sensors for Drug Development in Pharmaceutical Industry

Junrui Di, Marta Karas and Vanja Vlajnic

Abstract

As clinical trials evolve with technological advancements, wearable sensors and digital health technologies (DHTs) have significantly enhanced data collection by providing continuous, near real-time measurements. Traditional methods, constrained by infrequent site visits and subjective measures, often result in sparse, low-resolution data that limits understanding of patient outcomes. The adoption of wearables in drug development has led to the growth of novel digital endpoints across multiple therapeutic areas, such as stride velocity in Duchenne Muscular Dystrophy and physical activity in heart failure. Regulatory bodies have issued guidance supporting the integration of DHTs, emphasizing objective endpoints. The US Food and Drug Administration's Digital Health Center of Excellence and guidelines on remote data acquisition exemplify this support. Additionally, frameworks such as the Digital Medicine Society's "V3+" standardize the validation of fit-for-purpose digital endpoints. Emerging analytical approaches for wearable sensor data, including functional data analysis and handling missing data, further bolster the utility of digital endpoints in clinical trials. Collectively, these advancements allow for a more comprehensive and nuanced understanding of patient health, improving both the precision and applicability of clinical trial outcomes. Ultimately, the integration of digital endpoints revolutionizes patient monitoring, enhancing drug development and regulatory decision-making.

Keywords: digital endpoints, digital health technologies, verification and validation, clinical meaningfulness, data integration, functional data analysis, regulatory framework

1. Introduction

As technology improves, clinical trials have broadened the range and types of data collected. Traditionally, data collection has been limited to specific time points during the trial when patients are required to visit the site. During these visits, various types of data are collected, including medical history, laboratory results, blood tests, general wet biomarker data, clinical outcome assessments (COAs), and patient-reported

outcomes (PROs). Given the nature of certain assessments, such as wet biomarker data, patients must physically be present at the site for testing in many cases. One limiting factor of such trials is that these visits are often infrequent throughout the course of the trial, thus leading to relatively sparse data points for each patient. This low resolution of data can pose further analytical challenges, limiting the understanding of changes happening within a patient when data is not collected. Another common issue inherent to data such as patient-reported outcomes is the increased variability due to a subjective measure, which can lead to difficulties in interpretation of data. The application of sensors and wearables provide a unique opportunity to address some of these limitations.

The general development of sensor technologies and their applications in the pharmaceutical industry have greatly expanded over the last decades, enhancing our ability to objectively monitor the health and safety of patients enrolled in clinical trials. From accelerometers measuring physical activity in patients with cardiovascular disease, continuous glucose monitors (CGM) measuring glucose in patients with diabetes, to electroencephalograms (EEGs) measuring electrical activity in patients with epilepsy, applications of sensors within the pharmaceutical industry have become a vital component of elucidating the underpinnings of disease and understanding the patient journey for a specific disease. Specifically, sensors have a unique utility in that they are often able to continuously monitor a patient and provide measurements in near *real-time* which is a departure from traditional data collected as part of a clinical trial [1].

Several regulatory agencies have recently provided feedback and guidance on the use of sensors and wearables in the pharmaceutical industry. A main utility of sensors is that they provide objective measurements and endpoints. For example, as stated by the European Medicines Agency (EMA) in its guideline on medicinal products for the treatment of insomnia, two complimentary types of trial data are necessary to demonstrate efficacy in a clinical development program: 1) subjective endpoints, such as self-reported outcomes and surveys, and 2) objective endpoints such as those coming from polysomnography sensor data. As further stated in the guidelines, while individual criteria for improvement may be important for clusters of patients, clear clinical improvement based on data from objective measurements such as multichannel polysomnography is necessary to demonstrate clinical relevance and efficacy [2]. In the heart failure space, the U.S. Food and Drug Administration (FDA), has provided guidance that the effectiveness of a heart failure drug can be demonstrated based on improvements in clinical symptoms such as edema and fatigue, and/or also in functioning, such as activities of daily living, exercising, etc. FDA further states that it will consider trials that utilize novel endpoints such as measures of functional capacity and daily activity, e.g. accelerometry data, as acceptable endpoints to a clinical trial [3].

In summary, traditional measures such as examining physical limitation through a six-minute walking test (6MWT) are important to identify such things as maximum physical capacity, however, data from sensors and wearables can provide a complimentary data stream that gives insight to a patient's physical functioning on a day-to-day level. As noted by many regulatory bodies, sponsors, and payers, it is important to capture these functional capacity measurements to better understand some of the factors that patients feel are most reflective of their everyday life. Taken collectively with traditional data, data from sensors and wearables can provide a more cohesive understanding of a patient, the patient journey, and improve our understanding of important changes that are happening to the patient throughout the course of a clinical trial.

2. Current landscapes of novel digital endpoints

With the wide adoption of wearables devices in drug development, we have witnessed the rapid growth of digital endpoints to quantify treatment effects. In this section, we list examples of such novel digital endpoints across various indications.

2.1 Stride velocity 95th centile for in ambulatory functions in Duchenne muscular dystrophy

Duchenne muscular dystrophy (DMD) is a rare, genetic-mutation-induced, neuromuscular disorder that causes skeletal and cardiac muscle to degenerate and leads to progressive muscle weakness and declining motor function in the pediatric population [4]. Assessments of motor functions in patients with DMD conventionally heavily rely on COAs such as 6MWT and the North Star Ambulatory Assessment (NSAA) [5, 6], which have been shown to be strongly impacted by factors such as patients' motive, capability to understand instructions, and level of fatigue, etc. [7]

DMD patients typically have compromised walking capability, and thus they take significantly fewer steps and spend less time at moderate and high step rates [8]. Therefore, it is of great interest to quantify walking gaits using wearable sensors, which is an active area of research. Specifically, Stride Velocity 95th Centile (SV95C) is a digital measure of peak ambulation performance that represents the gait speed of the top 5% most rapid strides taken during a day in a real-world setting [7].

In 2019, SV95C measured by an ankle-worn wearable inertial sensor (ActiMyo®) was qualified by the EMA for use as a secondary endpoint in pivotal studies for DMD [9]. Furthermore, in 2023, it was qualified by EMA as the first approved primary endpoints for DMD in patients who are 4 years and older [10, 11]. It is currently being reviewed by FDA as a validated COA for DMD [12].

EMA particularly highlighted the advantages of using wearable-measured SV95C, such as, a continuous monitoring over relatively long periods in a home-setting, lower sensitivity to timing of the assessment (e.g. day and time of test), and less dependence on patient motivation or subjective assessment as compared to established tests. Other advantages demonstrated in the SPITFIRE/WN40227 trial of taldefgrobep alfa include sensitivity to ambulatory decline over short intervals, low variability, and correlation with established COAs [7].

2.2 Cardiovascular disease and heart failure

An example of a therapeutic area where sensors and wearable devices have had a strong impact in the drug development process is cardiovascular disease and heart failure. Traditionally, the 6MWT has been considered the gold standard clinical endpoint for assessing physical limitation in these patients. For the assessment, patients are tasked with walking as far of a distance as they can within six-minutes and the value is recorded. While assessing a patient's physical capability in a clinical setting provides valuable insights into their abilities, it only offers a limited view of their overall functioning. For instance, the 6MWT is conducted in a controlled environment, which may not accurately reflect the patient's real-life physical capabilities. In contrast, data collected from wearables and biosensors allows researchers to gain a deeper understanding of a patient's daily activity patterns in their natural surroundings. This type of data can help answer important questions, such as: 1) Over time, does the patient show an increase or decrease in daily physical activity levels? 2) Is the

patient able to sustain activity for longer durations? 3) What level of physical activity does the patient feel comfortable engaging in outside of the clinic? [13].

2.3 Wearable-measure nocturnal scratch for atopic dermatitis

Pruitus (itch) is the most common symptom for dermatological disorder such as atopic dermatitis (AtD), which notably is more severe during sleep and can impact sleep quality [14, 15]. Conventional clinical endpoints for AtD include clinician-administered COAs such investigator static global assessment (ISGA) [16] which examines the body surface area of lesion, and patient reported outcomes (PROs) such as peak pruitus numeric rating scale (PPNRS) [17]. COAs cannot provide a holistic picture of how a patients' life is impacted by skin conditions outside clinic, while PROs are subjective and may have recall biases [18].

Since the common reaction to the pruritus is scratching the affected area, which leads to additional inflammation and may further perpetuate the itch–scratch cycle, wearable sensors can be used to objectively quantify the scratching behavior [19]. In 2020, Yang et al. conducted a comprehensive literature review and identified 24 objective scratch measurements approaches using different types of sensors [20].

Most of these studies predominantly use accelerometers to capture specific wrist and finger movement patterns during scratching. A notable example is by Mahadevan et al. who developed the first accelerometer data-based solution to detect nocturnal sleep window and then predict scratch counts and durations within these windows, using a combination of heuristic signal processing algorithms and machine learning [21]. This method has been refined to incorporate more advanced signal features and machine learning models [22].

Other sensors that have been explored include acoustic, vibratory, strain gauge and pressure sensors. Though, research on such sensors is still limited and remains to be further evaluated in a larger setting. However, there are more user-friendly form factors of devices that have a multi-modal sensor built in which shows good potential. For example, a ring equipped with an accelerometer and a contact microphone has been shown to not only detect scratching episodes but also assess scratch intensity [23].

Another recent literature review took a further step to address the gap in terms of lacking standardization of the measure's definition and contextualization of nocturnal sleeping in order to bring forth a unified measurement definition for nocturnal scratch. Due to high prevalence of AtD and large market for its treatment, there has been cross-industry efforts to work in task forces to bring objectively measured nocturnal scratch to regulatory approval [24].

2.4 Polysomnography-derived measures of sleep and daytime sleepiness for narcolepsy

Narcolepsy type 1 (NT1) is a rare condition caused by a deficiency in orexin peptide. This deficiency disrupts the normal stimulation of brain regions responsible for wakefulness. Excessive daytime sleepiness (EDS) is a hallmark symptom of NT1, causing significant impairment in daily functioning and activities of patients [25–27].

Contemporary clinical trials of wake-promoting agents in NT1 primarily use the Maintenance of Wakefulness Test (MWT) and a self-reported Epworth Sleepiness Scale (ESS) to assess treatment efficacy for EDS [28–30]. The MWT is an in-lab test that evaluates sleep onset latency (SOL) under sleep-inducing conditions during patient's usual awake time. It typically consists of four 40-minute wake trials,

recorded with standard polysomnography (PSG) that includes electrooculography (EOG), chin electromyography (EMG), EEG, and electrocardiogram (EKG). The SOL is determined based on a sleep technician's assessment of PSG data, and a mean SOL across four wake trials is often used as an efficacy endpoint.

The rich PSG data from the MWT offers opportunities to derive and explore new objectively measured outcomes to characterize the agent's therapeutic effect on daytime wakefulness. A notable example is Tracey et al. [31] who examined "micro-sleeps" (sleep episodes lasting 3 to 15 seconds), sleep probability measures derived from automated sleep scoring, and EEG power band measures guided by previously reported correlations with subjective sleepiness [32]. These outcomes were evaluated in a phase 1B trial in NT1 and narcolepsy type 2, demonstrating response to treatment. Authors hypothesized that the novel EEG-based biomarkers allowed to capture an individual's ability to stay "truly awake and alert" (as opposed to "awake but drowsy"), complementing the traditionally used MWT outcome, SOL.

Characterizing the NT1 phenotype through nocturnal PSG (nPSG) is an active research area. Several sleep characteristics that distinguish narcolepsy patients from healthy controls have been identified, such as decreased sleep latency, REM sleep latency, and sleep efficiency [33–37]. While existing works have primarily focused on whole-night nPSG features, Vilela et al. [38] recently analyzed sleep in per quarter-night segments and found that the differences between NT1 and clinical controls were most pronounced in the first quarter of the night. nPSG-derived features could potentially be used as exploratory endpoints in clinical trials to assess the restoration of a healthy sleep phenotype in NT1. Additionally, a growing body of research explores the possibility of diagnosing narcolepsy using nPSG data alone [34, 39].

Finally, emerging wearable technologies for at-home EEG monitoring, [40] coupled with automated sleep staging software, may enable a scalable collection of longitudinal sleep measures in a naturalistic environment. These data may better reflect the daily burden of living with narcolepsy and other sleep-wake disorders than an episodical in-lab PSG visit. First examples of collaborations between at-home EEG biotech and pharmaceutical companies to support clinical trials have been announced [41].

2.5 Beyond efficacy endpoints: Other uses of digital wearable sensors in drug development

The examples above illustrate how wearable sensor data can be used to develop novel efficacy endpoints. Additionally, wearable sensors contribute to other drug development aspects, such as safety endpoints and patient diagnosis at screening, as shown below.

An ambulatory blood pressure monitor (ABPM) is a portable device for continuous blood pressure and heart rate measurement over 24 hours or more. It consists of a small monitor, worn in a pouch on a belt or under a pillow at night, and an arm cuff that inflates during measurements [42]. ABPM is commonly used in clinical practice to diagnose hypertensive disorders [42]. In hypertension treatment development, blood pressure outcomes assessed through ABPM may serve as a primary endpoint (NCT00591578). ABPM may also be used for trial screening (e.g., to rule in/out a hypertensive disorder) and in other indications for safety evaluations. Specifically, in 2022, the FDA issued a revised draft guidance, "Assessment of Pressor Effects of Drugs Guidance for Industry" [43]. This guidance advises sponsors on assessing unintended blood pressure effects in the context of all drugs in development. The

FDA recommends including in a trial a thorough blood pressure assessment for any drug intended for chronic use, and recommends using ABPM for this purpose, as it can detect small but potentially relevant changes in blood pressure.

The Home Sleep Apnea Test (HSAT) device is a portable tool used to diagnose obstructive sleep apnea (OSA), a condition characterized by pauses in breathing during sleep. It is designed to enable unsupervised use at home. The HSAT device typically consists of a small recorder worn on a belt around the chest, a nasal cannula with sensor prongs to be placed inside nostrils, and a finger sensor to measure oxygen saturation. These sensors connect to the recorder, which tracks various measures, including breathing effort, body position, air pressure during inhalation and exhalation, snoring, pulse rate, and oxygen saturation [44]. In clinical trials in people with moderate to severe OSA, the HSAT device has been used as an alternative to in-lab PSG for moderate or severe OSA diagnosis during screening (NCT05814016). Notably, in the "Clinical Use of a Home Sleep Apnea Test: An Updated American Academy of Sleep Medicine Position Statement" in 2018, [45] the American Academy of Sleep Medicine (AASM) recommends that HSATs may be used as an alternative to PSG for diagnosing OSA in uncomplicated adults with signs of moderate to severe OSA, but only after a face-to-face examination by a medical provider (in person or via

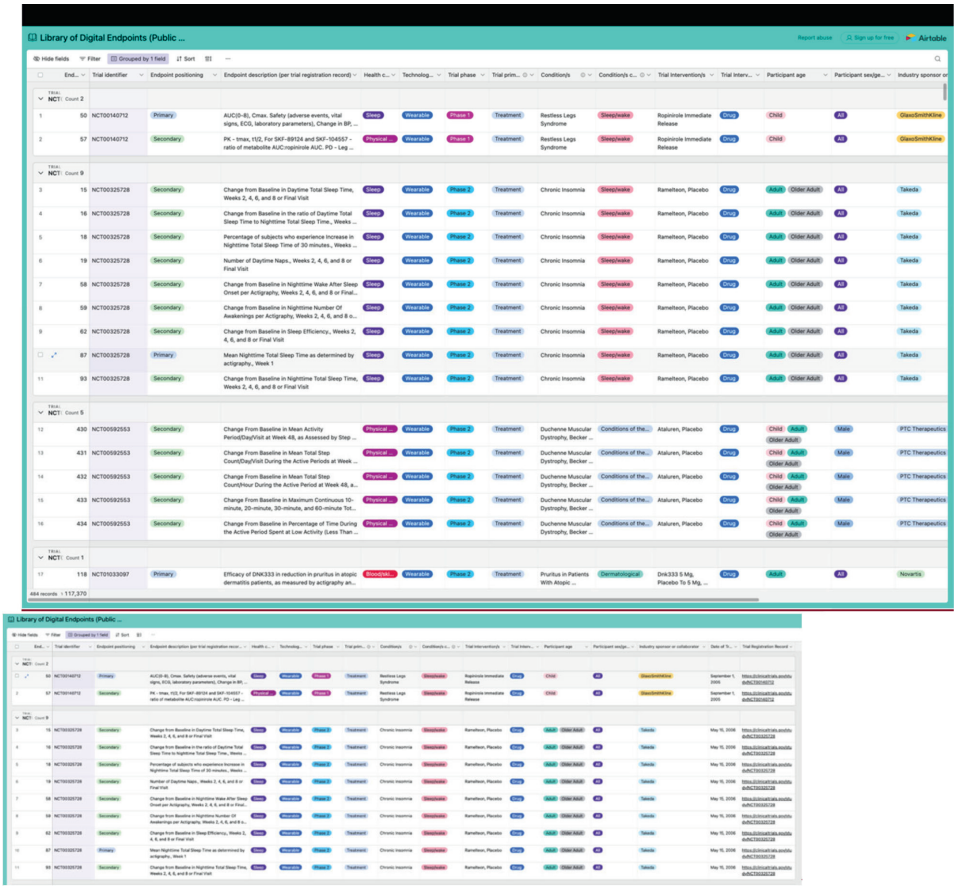

Figure 1.
Digital endpoints library.

telemedicine). The diagnosis "must not be based solely on automatically scored HSAT data" and the raw data from HSAT devices must be reviewed by a sleep medicine specialist.

2.6 Digital medicine society (DiME)'s digital endpoints library

Digital Medicine Society (DiME) was founded in 2020 as a professional society for the advancement of the digital medicine community. It has since inspired multiple pre-competitive communities and consortiums and moreover established multiple cross-industry frameworks to ensure the proper development and application of digital endpoints.

As a rapidly growing field, there have been continuous efforts to deploy wearable sensors and use them to collect clinical endpoints for trials. In order to keep up with the pace, DiME maintains a live resource, i.e. the digital endpoints library, benchmarks progress in the field and highlights the work we must continue to do to advance the use of digital endpoints to speed medical product development. As of mid 2024, the link to the library is https://dimesociety.org/library-of-digital-endpoints/. The library is organized as a dashboard displaying trial-related information such as 1) sponsors/times/NCT registration number, 2) endpoints used and whether it is primary or secondary, 3) clinical indication, 4) phase of the study, 5) symptoms of interest, etc. This library serves as a resource for industry sponsors to leverage what has been done in the field. A screenshot of the library can be seen in **Figure 1**.

3. Regulatory and industry-wise guidance and framework on fit-for-purpose use of wearable sensors

The successful adoption of wearable sensors in drug development requires joint efforts from across the industry, including sponsors, device companies, research institutes, and regulators etc. In order to summarize the lessons learned and avoid reinventing the wheel, there has been cross-industry efforts in developing generalizable guidance and framework for the development and application of fit-for-purpose digital endpoints.

3.1 Regulatory efforts and guidance

The FDA has publicly stated the potential benefits digital health technologies (DHTs) such as wearables bring to the drug development process and is committed to supporting the use of DHTs in clinical drug development. Meanwhile, FDA has developed a comprehensive program to collaborate with relevant parties in this space. In September 2020, FDA launched the Digital Health Center of Excellence to empower stakeholders and advance health care by fostering responsible and high-quality digital health innovation [46]. This center created a diverse series of programs and frameworks to promote the wide use of DHTs.

Internally, FDA has established the DHT Steering Committee which consists of senior staff from several of their internal divisions, and other branches to oversee activities to assist organizational units with consistent approaches to the review of drug submissions that contain DHT-derived data [47, 48]. The new group provided technical expertise and trainings in key aspects such as endpoint verification and validation, device updates, and the use of a participant's own DHT [47].

FDA is also proactively engaged with external stakeholders through various programs such as the Drug Development Tool (DDT) Qualification Program [49] which aims to support the development of tools for use in assessing medical products. The program focuses on developing tools that may be relied upon in multiple clinical investigations to support premarket submissions for drugs where the context of use is the same.

The other major effort from FDA is the release of the final guidance on "Digital Health Technologies for Remote Data Acquisition in Clinical Investigations" which provides recommendations on the appropriate use of DHTs in clinical trials [1]. As the first of its kind official regulatory guidance, it provides consensus and expectation on a couple of key aspects as illustrated in **Figure 2**. For study design, operation, and submission with DHTs, the sponsor has to ensure and document the device is fit-for-purpose. For device verification and validation, the document sets up expectations, such as type of validations against reference measurements. It also specifies the potential difference between remote and site-based clinical data collection, and that certain validation studies and usability evaluations can be conducted in healthy volunteers and individuals with varying degrees of disease severity. The guidance also describes what defines and constitutes a novel digital endpoint that captures the concept of interest. Furthermore, it states the importance of considering events that may impact data collection and interpretation, and missing data under the estimand framework. Last but not least, it emphasizes the specific risk associated with clinical events and device-related privacy and how informed consent should be used to appropriately inform the patients on how their device data will be collected and used to ensure protection of data and trust.

The EMA has also released several pieces of guidance related to wearables and sensors [50–52]. The "Regulatory Science Strategy to 2025", developed by the EMA outlines strategic goals directly relevant to DHTs [53]. Additionally, the EMA has actively explored the use of digital measures in specific drug development contexts, such as the SV95C for DMD discussed earlier.

3.2 Digital medicine society (DiME)'s efforts and guidance

DiME's hallmark contribution to the field is the so-called "V3" framework released in 2020, as the first of its kind cross-industry framework to evaluate the developments of fit-for-purpose digital endpoints using digital health technologies [54]. This framework was further enhanced to "V3+" which defines the scopes and responsible

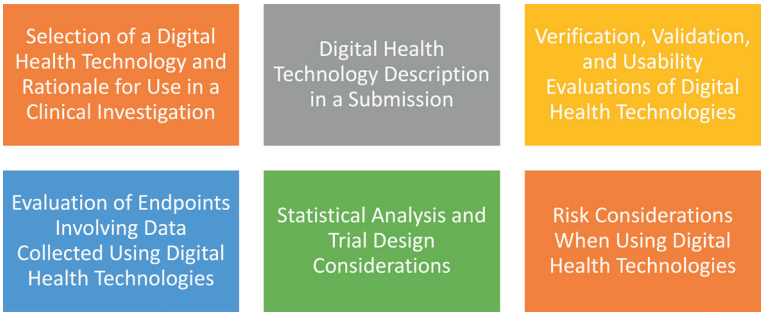

Figure 2.
FDA's DHT guidance.

parties for 1) verification, 2) usability validation, 3) analytical validation, and 4) clinical validation before any clinical utility, as illustrated in **Figure 3** [55]. The framework clarifies core terminology and best practices for the evaluation of novel digital endpoints and the device they are collected from. The "who, what, when, where, why", together with the expected documentations and clinical rational for each component has been defined. Specifically, the added "usability validation" aims to ensure user-centricity and scalability of high-quality digital measurement tools, and pre-identified potential user-related risks. Since its establishment, "V3" (and inherently "V3+") have been widely utilized for deploying a new digital endpoint in clinical trials. However, it is important to note that the framework has not been formally endorsed by regulators.

Another of DiME's contribution is the release of 'The Playbook – Digital Clinical Measures" to guide remote monitoring across clinical research, clinical care, and public health [56]. The Playbook defined a foundation for developing and deploying digital clinical measures based on three components, namely, 1) *measure*, i.e. what do you want to measure and why; 2) *technology*, i.e. what are the right tools for the job, and 3) *operation,* i.e. what's needed to deploy remotely at scale. For example, as for the *measure* component, the Playbook proposes the 3-step procedure to develop meaningful digital measures: determine the meaningful aspect of health (e.g. "I want to be more physical active"), identify the concept of interest (e.g. "physical activity"), and identify outcomes to be measured (e.g. "daily step counts"). Additionally, the "V3+" framework provides a validation procedure to choose the "right technology". At last, key operational considerations may include, device distribution, site training, and compliance monitoring etc. In summary, the framework offers a checklist that covers several aspects of clinical studies with wearable sensors and other digital health technologies.

3.3 Clinical validity and clinical meaningfulness

As suggested by the FDA's guidance and DiME's "V3+" framework, clinical validation is a key step to ensure the use of device-measured digital endpoints which can be clinically utilized to develop new medical product in patients with a specific indication and can answer the clinical question of interest. It evaluates whether the device acceptably identifies, measures, or predicts a meaningful clinical, biological, physical, functional state, or experience in the specified context of use [54].

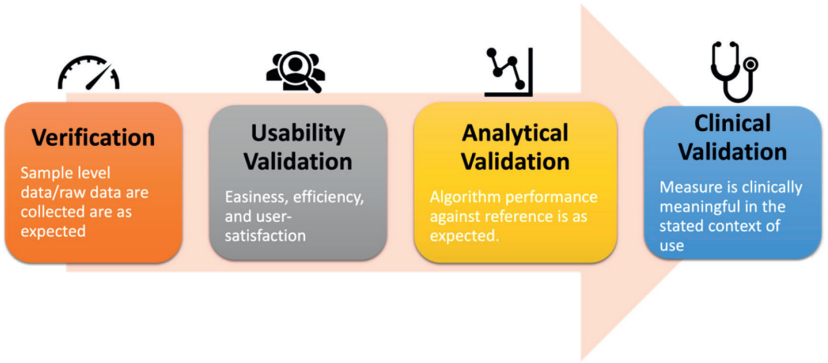

Figure 3.
DiME's V3+ framework.

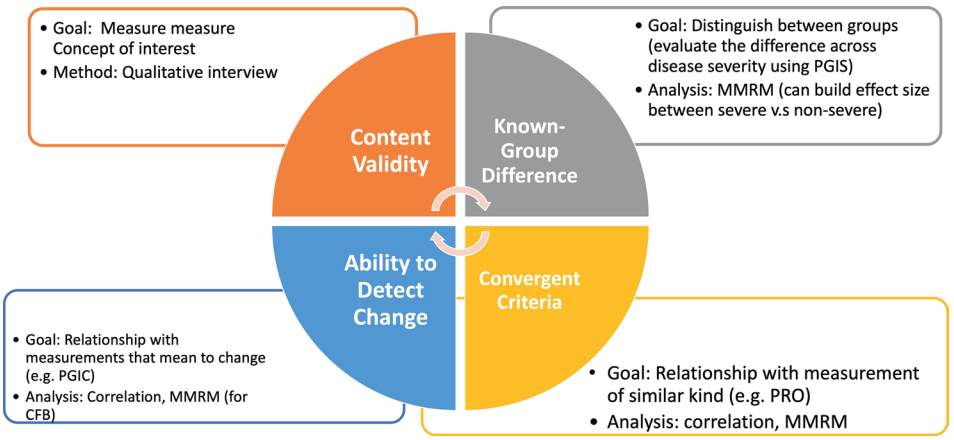

Figure 4.
Conceptual and analytical framework for clinical validation. PGIS: patient global impression of severity. PGIC: patient global impression of change. MMRM: mixed models with repeated measurement. CFB: change from baseline. PRO: patient reported outcomes.

The recent EMA opinion document of SV95C for DMD as the primary endpoints specifies a framework and the associated analytical considerations to evaluate the clinical validity of digital endpoints for a specific indication [11]. It is worth noticing that such a framework was adopted from psychometric analyses for the validation of PROs, which covers the psychometric properties of PROs such as content validity, construct validity, and responsiveness (or the ability to detect change) [57]. The key components in such framework and the commonly used analytical tools are summarized in **Figure 4**. The evaluation of "known group difference" and "convergent criteria", together define the construct validity which means the degree to which the measurement of interest are consistent with internal relationships, relationships with scores of other instruments, or differences between relevant groups [58].

Such framework will also set up the foundation for the next step to determine within-patient meaningful change (WPMC, i.e. improvement and deterioration from the patients' perspective) using either anchor-based methods or distribution-based methods, or both [59]. The FDA's Patient Focus Drug Development Guidance 4 specifically laid the foundation for the anchor-based and distribution-based methods, such as how sponsors should select the recall periods and anchors used, and the fact that how one should supplement anchor-based methods with empirical distribution-based methods [60].

4. Emerging analytical approaches for clinical trials with wearable sensors

Many wearable sensors such as accelerometers and CGM, collect data passively and frequently, leading to rich streams of complex time-series signals from which outcome measures are derived using algorithms. Careful consideration for analytical approaches is needed when deriving clinical endpoints in drug development programs to ensure that they are robust and reliable to enable regulatory decision making. In this section, we review emerging statistical analysis methods for wearable sensor-based data.

4.1 Integration of wearable-measured data from multiple domains or modalities

Nowadays, a number of sensors (e.g. heart rate monitor, thermometer, pulse oximeter (SpO2), continuous glucose monitoring) can be packaged together in one integrated wearable device (e.g. Apple Watch). In the meantime, even the same device can simultaneously measure a wide variety of different physiological domains. For instance, a single accelerometer can measure day-time physical activity, night time sleep quantity/quality, and 24-hour circadian rest/activity rhythms, each has its own clinical implication [61]. With features derived from multiple domains or modalities, one can directly combine those features in linear or nonlinear fashions using statistical and machine learning models [62].

However, another innovative way to evaluate data from multiple modalities is to separate the joint effects that are homogenous across different modalities from the modal-specific effects. The Joint and Individual Variation Explained (JIVE) approach was developed to serve this purpose, and can decompose multi-modal data into a low-rank approximation capturing joint variation across data types, low-rank approximations for structured variation individual to each data modal [63]. JIVE was applied to integrate accelerometry-derived features quantifying three physiological domains of activity, sleep, and circadian rhythm to quantify and separate between- and within-domain variation [64]. More recently, JIVE has been applied to evaluate the associations between joint and individual components from activity, sleep, and circadian rhythm with current and remitted major depression disorder [65]. The same concept can be directly implemented to study features derived from multi-modal sensors and such an approach can be further evaluated for clinical use.

4.2 Functional data analysis

Modern sensors can measure various physiological variables in a near-continuous manner. This provides an opportunity to have a granular assessment of individual's and population's trends. For example, trends across time of a day can be studied in physical activity data recorded at minute-level across a day(s). Such analyses can be handled with functional data (FD) analysis, a relatively new field of statistics that deals with data where an ideal unit of observation is a function defined on some continuous domain (as opposed to a scalar value), and the observed data is a function recorded on a discrete grid [66].

On the technical side, many FD analysis methods approximate a vector of observations through a basis of functions, offering several advantages over traditional multivariate techniques that could be alternatively considered. FD analysis allows for the representation of information in low-dimensional spaces, possibly leading to more powerful hypothesis testing and better predictive models. Additionally, FD analysis methods may reduce noise and outliers' effect, improve computational efficiency, and enable evaluation of the underlying function at any point in a continuous domain (as opposed to being restricted to a set of points of data collection), offering better interpretability [67, 68].

Several FD analysis techniques and models have been proposed in recent years. Functional principal component analysis, [69] an extension of multivariate principal components analysis, describes the variability of a sample of curves by decomposing the space of curves into principal directions of variation. Functional regression models incorporate functional outcomes and/or functional predictors into the model [70–72]. For instance, in function-on-scalar regression, a functional coefficient

represents a varying nonlinear effect of a covariate on the functional outcome. Recent advancements have extended generalized linear models and other frameworks to include functional components and address multilevel functional data structures [73–76].

On the application side, FD analysis has been widely used to quantify physical activity from wearable accelerometer data, [77–79] glucose profiles from continuous glucose monitors, [80] white-matter tract profiles from diffusion tensor imaging, [81] cellular organization in tumor microenvironment studies, [82] and changes in sleep-related arousal after ketamine treatment in people with depression [83]. In drug development, FD regression has been applied to model dose–response curves [84]. As FD techniques advance in both theory and availability of methods implementation, [85, 86] they are expected to play a growing role in developing digital biomarkers, enabling a more detailed quantification of patterns in wearable sensor data.

4.3 Composite endpoints

Composite endpoints combine multiple clinically relevant outcomes into a single measure that may be used in clinical trials to assess treatment efficacy, particularly when diseases have complex, multidimensional presentations. They may help improve the statistical power of the tests, allowing for smaller sample sizes and shorter follow-up periods, but they require careful interpretation to avoid misinforming clinicians and policymakers [87, 88]. Since wearable sensors can simultaneously measure multiple physiological variables, it is natural to consider creating composite endpoints by combining several digital measures into a single outcome.

Notable examples come from monitoring disease progression in amyotrophic lateral sclerosis (ALS), a neurodegenerative disease characterized by a gradual loss of motor neurons, leading to progressive muscle weakness and eventually death [89]. To date, a few works have demonstrated the use of individual measures derived from wearable accelerometers, step-counters, and smartphone sensors to quantify ALS progression [90–93]. In Gupta and colleagues, [94] a composite endpoint was created by combining measures derived from accelerometer data collected continuously at home to produce severity scores for each limb via machine-learning. The authors showed that the obtained severity scores demonstrated faster progression (larger estimated decline per year with smaller standard deviation) than ALS Functional Rating Scale-Revised (ALSFRS-R), a commonly used scale to track ALS progression in contemporary clinical trials. Neumann and colleagues [95] developed a composite endpoint by creating optimal linear combination of multiple speech biomarkers to enhance discrimination between bulbar and non-bulbar onset in ALS and disease progression monitoring. The authors emphasized the clinical interpretability of their composite endpoint due to the understanding of relative contributions of each feature, making their approach potentially valuable for ALS pharmaceutical trials.

4.4 Address missing data from wearables

While conventionally much work has been reported on the evidence to support device selection and reliability of digital measures, in recent years, there has been consensus on the importance of how to deal with missing data in wearable sensor-collected time-series data.

In 2022, Di et al. provided discussion on handling missing data from the perspective of study design, operation, and analytic approaches [96]. The work reviewed characteristics of wearable sensor data that cause issues related to missing device data. For those researchers who are considering analyzing wearable sensor-based data, this work provided a reference for statistically sound approaches to address missingness, and can inspire further methodological innovation and applications to correctly handle DHT missing data, enabling the derivation of reliable and clinically meaningful digital endpoints for use in clinical trials. This paper presented some techniques such as sparse functional data analysis to analyze continuous time-series signals with a chunk of missing data in specific time windows, and treating partially observed days (i.e. days with shorter than enough wearing time) as right-censored data.

4.5 Artificial intelligence (AI) and machine learning

The vast variety, large volumes, and immense potential of data collected by wearable devices make this domain an ideal candidate for leveraging advanced AI/ML methods. In turn, the continuous, high-dimensional data from wearables helps fuel AI/ML techniques to create novel DHT-based data products.

One example of an ML-powered product is the development of nocturnal scratch detection discussed above, which demonstrates an ideal use case for supervised ML methods. The continuous nature of device signals is perfectly suited to deep learning models like convolutional neural networks (CNNs) and recurrent neural networks (RNNs). For instance, Acharya and Basu developed a hybrid CNN-RNN model to classify breathing sound anomalies (e.g., wheezes, crackles) for the automated diagnosis of respiratory and pulmonary diseases using various recording devices [97]. Similarly, Hssayeni et al. applied a long short-term memory (LSTM) model to detect tremor severity (both resting and action tremors) in Parkinson's disease (PD) patients using accelerometer and gyroscope data [98].

However, labeling large amounts of data is both costly and time-consuming. As a result, methods that efficiently exploit unlabeled data while minimizing labeling efforts are highly desirable. Semi-supervised learning approaches are particularly promising in this context, as they can utilize a combination of limited labeled data and large volumes of unlabeled data. For example, Yang et al. developed a semi-supervised learning algorithm that identifies near-miss falls among ironworkers without disrupting jobsite activities, potentially preventing fall accidents through the use of wearable inertial measurement units [99].

5. Concluding remark

The integration of novel digital wearable sensors into drug development holds significant promise for transforming how pharmaceutical companies measure clinical endpoints across a range of indications. As discussed, wearables offer a unique ability to capture continuous, real-world data, providing deeper insights into patient health and treatment efficacy. The FDA's guidance on DHT and frameworks like DiME's "V3" are crucial in standardizing the use of these technologies, ensuring their adoption is both scientifically rigorous and clinically meaningful.

While the potential is vast, several challenges remain. The primary hurdles include ensuring data accuracy and reliability, addressing privacy concerns, increasing patients' compliance, and overcoming the complexities of data integration

into existing regulatory pathways. Moreover, the need for innovative analytical approaches to make sense of the vast amounts of data generated by wearables cannot be overstated. As the industry continues to navigate these challenges, collaboration between technology developers, regulatory bodies, and pharmaceutical companies will be essential.

Ultimately, the future of drug development may well hinge on how effectively these challenges are addressed. Wearable sensors represent not just a technological advancement but a paradigm shift in how we approach patient monitoring and treatment development, promising a more personalized and precise era of medicine.

Conflict of interest

Junrui Di is an employee of Dexcom, Inc. Marta Karas is an employee of Takeda Development Center Americas, Inc., and a stockholder in Takeda Pharmaceutical Company Limited. Vanja Vlajnic is an employee of Bayer US.

Author details

Junrui Di[1], Marta Karas[2] and Vanja Vlajnic[3]*

1 Clinical Affairs, Dexcom, Inc., San Diego, CA, USA

2 Statistical and Quantitative Sciences, Takeda Development Center Americas, Inc., Cambridge, MA, USA

3 Clinical Statistics and Analytics, Bayer US, Whippany, NJ, USA

*Address all correspondence to: vanja.vlajnic@bayer.com

IntechOpen

References

[1] US FDA. Digital Health Technologies for Remote Data Acquisition in Clinical Investigations [Internet]. U.S. Food and Drug Administration. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/digital-health-technologies-remote-data-acquisition-clinical-investigations; 2021 [Accessed: September 8, 2024]

[2] European Medicines Agency. Guideline on Medicinal Products for the Treatment of Insomnia [Internet]. European Medicines Agency; 2011. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-medicinal-products-treatment-insomnia-revision-1_en.pdf [Accessed: September 15, 2024]

[3] US FDA. Treatment for Heart Failure: Endpoints for Drug Development Guidance for Industry [Internet]. 2019. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/treatment-heart-failure-endpoints-drug-development-guidance-industry [Accessed: September 8, 2024]

[4] Falzarano M, Scotton C, Passarelli C, Ferlini A. Duchenne muscular dystrophy: From diagnosis to therapy. Molecules. 2015;**20**(10):18168-18184

[5] McDonald CM, Henricson EK, Han JJ, Abresch RT, Nicorici A, Elfring GL, et al. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. Muscle & Nerve. 2010;**41**(4):500-510

[6] Scott E, Eagle M, Mayhew A, Freeman J, Main M, Sheehan J, et al. Development of a functional assessment scale for ambulatory boys with Duchenne muscular dystrophy. Physiotherapy Research International. 2012;**17**(2):101-109

[7] Rabbia M, Guridi Ormazabal M, Staunton H, Veenstra K, Eggenspieler D, Annoussamy M, et al. Stride velocity 95th centile detects decline in ambulatory function over shorter intervals than the 6-minute walk test or north star ambulatory assessment in Duchenne muscular dystrophy. Journal of Neuromuscular Diseases. 2024;**11**(3):701-714

[8] McDonald CM, Widman LM, Walsh DD, Walsh SA, Abresch RT. Use of step activity monitoring for continuous physical activity assessment in boys with Duchenne muscular dystrophy. Archives of Physical Medicine and Rehabilitation. 2005;**86**(4):802-808

[9] European Medicines Agency. Qualification Opinion on Stride Velocity 95th Centile as a Secondary Endpoint in Duchenne Muscular Dystrophy Measured by a Valid and Suitable Wearable Device* [Internet]. European Medicines Agency; 2019. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-stride-velocity-95th-centile-secondary-endpoint-duchenne-muscular-dystrophy_en.pdf [Accessed: September 8, 2024]

[10] Servais L, Eggenspieler D, Poleur M, Grelet M, Muntoni F, Strijbos P, et al. First regulatory qualification of a digital primary endpoint to measure treatment efficacy in DMD. Nature Medicine. 2023;**29**(10):2391-2392

[11] European Medicines Agency. Draft Qualification Opinion for Stride Velocity 95th Centile as Primary

Endpoint in Studies in Ambulatory Duchenne Muscular Dystrophy Studies [Internet]. European Medicines Agency; 2023. Available from: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/draft-qualification-opinion-stride-velocity-95th-centile-primary-endpoint-studies-ambulatory-duchenne-muscular-dystrophy_en.pdf [Accessed: September 8, 2024]

[12] US FDA. DDT COA #000103: ActiMyo® [Internet]. 2018. Available from: https://www.fda.gov/drugs/clinical-outcome-assessment-coa-qualification-program/ddt-coa-000103-actimyor [Accessed: September 8, 2024]

[13] Butler J, Khan MS, Gasior T, Erickson TR, Vlajnic V, Kramer F, et al. Accelerometer vs. other activity measures in heart failure with preserved ejection fraction: The VITALITY-HFpEF trial. ESC Heart Failure. 2024;**11**(1):293-298

[14] Yosipovitch G, Rosen JD, Hashimoto T. Itch: From mechanism to (novel) therapeutic approaches. Journal of Allergy and Clinical Immunology. 2018;**142**(5):1375-1390

[15] Lavery M, Stull C, Kinney M, Yosipovitch G. Nocturnal pruritus: The battle for a peaceful night's sleep. International Journal of Molecular Sciences. 2016;**17**(3):425

[16] Thyssen JP, Zang C, Neary MP, Bushmakin AG, Cappelleri JC, Cha A, et al. Translating the investigator's static global assessment to the eczema area and severity index in studies of crisaborole for atopic dermatitis. Dermatology and Therapy (Heidelb). 2021;**11**(3):845-853

[17] Yosipovitch G, Reaney M, Mastey V, Eckert L, Abbé A, Nelson L, et al. Peak pruritus numerical rating scale: Psychometric validation and responder definition for assessing itch in moderate-to-severe atopic dermatitis. British Journal of Dermatology. 2019;**181**(4):761-769

[18] Murray C, Rees J. Are subjective accounts of itch to be relied on? The lack of relation between visual analogue itch scores and actigraphic measures of scratch. Acta Dermato Venereologica. 2011;**91**(1):18-23

[19] Ebata T, Aizawa H, Kamide R. An infrared video camera system to observe nocturnal scratching in atopic dermatitis patients. The Journal of Dermatology. 1996;**23**(3):153-155

[20] Yang AF, Nguyen M, Li AW, Lee B, Chun KS, Wu E, et al. Use of technology for the objective evaluation of scratching behavior: A systematic review. JAAD International. 2021;**5**:19-32

[21] Mahadevan N, Christakis Y, Di J, Bruno J, Zhang Y, Dorsey ER, et al. Development of digital measures for nighttime scratch and sleep using wrist-worn wearable devices. NPJ Digital Medicine [Internet]. 2021;**4**(1):42. Available from: http://www.nature.com/articles/s41746-021-00402-x

[22] Ji J, Venderley J, Zhang H, Lei M, Ruan G, Patel N, et al. Assessing nocturnal scratch with actigraphy in atopic dermatitis patients. NPJ Digital Medicine. 2023;**6**(1):72

[23] Padmanabha A, Choudhary S, Majidi C, Erickson Z. A multimodal sensing ring for quantification of scratch intensity. Communications Medicine. 2023;**3**(1):115

[24] Ke Wang W, Cesnakova L, Goldsack JC, Dunn J. Defining the digital measurement of scratching during sleep or nocturnal scratching: Review of the literature. Journal of Medical Internet Research. 2023;**25**:e43617

[25] Dauvilliers Y, Arnulf I, Mignot E. Narcolepsy with cataplexy. The Lancet [Internet]. 2007;**369**(9560):499-511. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(07)60237-2/fulltext

[26] Plazzi G, Serra L, Ferri R. Nocturnal aspects of narcolepsy with cataplexy. Sleep Medicine Reviews [Internet]. 2008;**12**(2):109-128. Available from: https://www.sciencedirect.com/science/article/pii/S1087079207001232

[27] Scammell TE. Narcolepsy. New England Journal of Medicine [Internet]. 2015;**373**(27):2654-2662. DOI: 10.1056/NEJMra1500587

[28] Johns MW. A new method for measuring daytime sleepiness: The Epworth sleepiness scale. Sleep [Internet]. 1991;**14**(6):540-545. DOI: 10.1093/sleep/14.6.540

[29] Mitler MM, Gujavarty KS, Browman CP. Maintenance of wakefulness test: A polysomnographic technique for evaluating treatment efficacy in patients with excessive somnolence. Electroencephalography and Clinical Neurophysiology [Internet]. 1982;**53**(6):658-661. Available from: https://www.sciencedirect.com/science/article/pii/0013469482901420

[30] Zhan S, Ye H, Li N, Zhang Y, Cheng Y, Wang Y, et al. Comparative efficacy and safety of multiple wake-promoting agents for the treatment of excessive daytime sleepiness in narcolepsy: A network meta-analysis. Nature and Science of Sleep. 2023;**15**:217-230

[31] Tracey B, Culp M, Fabregas S, Mignot E, Buhl DL, Volfson D. Novel biomarkers derived from the maintenance of wakefulness test as predictors of sleepiness and response to treatment. Sleep [Internet]. 2024:zsae148. DOI: 10.1093/sleep/zsae148

[32] Strijkstra AM, Beersma DGM, Drayer B, Halbesma N, Daan S. Subjective sleepiness correlates negatively with global alpha (8-12 Hz) and positively with central frontal theta (4-8 Hz) frequencies in the human resting awake electroencephalogram. Neuroscience Letters [Internet]. 2003;**340**(1):17-20. Available from: https://www.sciencedirect.com/science/article/pii/S0304394003000338

[33] Barateau L, Lopez R, Chenini S, Rassu AL, Scholz S, Lotierzo M, et al. Association of CSF orexin-A levels and nocturnal sleep stability in patients with hypersomnolence. Neurology. 2020;**95**(21):e2900-e2911

[34] Christensen JAE, Carrillo O, Leary EB, Peppard PE, Young T, Sorensen HBD, et al. Sleep-stage transitions during polysomnographic recordings as diagnostic features of type 1 narcolepsy. Sleep Medicine [Internet]. 2015;**16**(12):1558-1566. Available from: https://www.sciencedirect.com/science/article/pii/S1389945715008242

[35] Hansen MH, Kornum BR, Jennum P. Sleep–wake stability in narcolepsy patients with normal, low and unmeasurable hypocretin levels. Sleep Medicine [Internet]. 2017;**34**:1-6. Available from: https://www.sciencedirect.com/science/article/pii/S1389945717301053

[36] Maski K, Mignot E, Plazzi G, Dauvilliers Y. Disrupted nighttime sleep and sleep instability in narcolepsy. Journal of Clinical Sleep Medicine. 2022;**18**(1):289-304

[37] Zhang Y, Ren R, Yang L, Zhang H, Shi Y, Sanford LD, et al. Polysomnographic nighttime features

of narcolepsy: A systematic review and meta-analysis. Sleep Medicine Reviews [Internet]. 2021;**58**:101488. Available from: https://www.sciencedirect.com/science/article/pii/S1087079221000733

[38] Vilela M, Tracey B, Volfson D, Barateau L, Cai A, Buhl DL, et al. Identifying time-resolved features of nocturnal sleep characteristics of narcolepsy using machine learning. Journal of Sleep Research [Internet]. 2024:e14216. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/jsr.14216

[39] Stephansen JB, Olesen AN, Olsen M, Ambati A, Leary EB, Moore HE, et al. Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. Nature Communications [Internet]. 2018;**9**(1):5229. Available from: https://www.nature.com/articles/s41467-018-07229-3

[40] Yoon H, Choi SH. Technologies for sleep monitoring at home: Wearables and nearables. Biomedical Engineering Letters [Internet]. 2023;**13**(3):313-327. DOI: 10.1007/s13534-023-00305-8

[41] Beacon Biosignals. Beacon Biosignals and Takeda Launch Strategic Collaboration to Advance Neurobiomarkers and Endpoints for Sleep Disorders [Internet]. Beacon Biosignals; 2024. Available from: https://beacon.bio/press-releases/beacon-biosignals-and-takeda-launch-strategic-collaboration-to-advance-neurobiomarkers-and-endpoints-for-sleep-disorders/ [Accessed: September 8, 2024]

[42] Pena-Hernandez C, Nugent K, Tuncel M. Twenty-four-hour ambulatory blood pressure monitoring. Journal of Primary Care & Community Health [Internet]. 2020;**11**. Available from: https://journals.sagepub.com/doi/full/10.1177/2150132720940519

[43] US FDA. Assessment of Pressor Effects of Drugs Guidance for Industry [Internet]. U.S. Food and Drug Administration; 2022. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/assessment-pressor-effects-drugs-guidance-industry

[44] Penzel T, Schöbel C, Fietze I. New technology to assess sleep apnea: Wearables, smartphones, and accessories. F1000Res [Internet]. 2018;**7**:413. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5883394/

[45] Rosen IM, Kirsch DB, Carden KA, Malhotra RK, Ramar K, Aurora RN, et al. Clinical use of a home sleep apnea test: An updated American Academy of sleep medicine position statement. Journal of Clinical Sleep Medicine [Internet]. 2018;**14**(12). Available from: https://jcsm.aasm.org/doi/10.5664/jcsm.7540

[46] US FDA. Digital Health Center of Excellence [Internet]. Available from: https://www.fda.gov/medical-devices/digital-health-center-excellence [Accessed: September 8, 2024]

[47] US FDA. Framework for the Use of Digital Health Technologies in Drug and Biological Product Development [Internet]. U.S. Food and Drug Administration; 2023. Available from: https://www.fda.gov/media/166396/download?attachment [Accessed: September 8, 2024]

[48] US FDA. Digital Health Technologies (DHTs) for Drug Development [Internet]. Available from: https://www.fda.gov/science-research/science-and-research-special-topics/digital-health-technologies-dhts-drug-

development#framework [Accessed: September 8, 2024]

[49] US FDA. Drug Development Tool (DDT) Qualification Programs [Internet]. Available from: https://www.fda.gov/drugs/development-approval-process-drugs/drug-development-tool-ddt-qualification-programs [Accessed: September 8, 2024]

[50] European Medicines Agency. Essential Considerations for Successful Qualification of Novel Methodologies [Internet]. European Medicines Agency; 2017. Available from: https://scrip.citeline.com/-/media/supporting-documents/pink-sheet/2017/12/ema-novel-methodology-qualification.pdf [Accessed: September 8, 2024]

[51] European Medicines Agency. Qualification Opinion on eSource Direct Data Capture (DDC) [Internet]. European Medicines Agency; 2019. Available from: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-esource-direct-data-capture-ddc_en.pdf [Accessed: September 8, 2024]

[52] European Medicines Agency. Questions and Answers: Qualification of Digital Technology-Based Methodologies to Support Approval of Medicinal Products [Internet]. European Medicines Agency; 2020. Available from: https://fnih.org/sites/default/files/2023-04/questions-answers-qualification-digital-technology-based-methodologies-support-approval-medicinal_en.pdf [Accessed: September 8, 2024]

[53] European Medicines Agency. Regulatory Science to 2025 [Internet]. European Medicines Agency; 2020. Available from: https://www.ema.europa.eu/en/annual-report-2019/regulatory-science-strategy-2025.html [Accessed: September 8, 2024]

[54] Goldsack JC, Coravos A, Bakker JP, Bent B, Dowling AV, Fitzer-Attas C, et al. Verification, analytical validation, and clinical validation (V3): The foundation of determining fit-for-purpose for biometric monitoring technologies (BioMeTs). NPJ Digital Medicine [Internet]. 2020;**3**(1):55. Available from: http://www.nature.com/articles/s41746-020-0260-4

[55] V3+: An Extension to the V3 Framework to Ensure User-Centricity and Scalability of Sensor-Based Digital Health Technologies [Internet]. Available from: https://dimesociety.org/courses/building-fit-for-purpose-sensor-based-digital-health-technologies/ [Accessed: September 8, 2024]

[56] Digital Medicine Society (DiME). The Playbook: Digital Clinical Measures [Internet]. Digital Medicine Society (DiME); 2023. Available from: https://playbook.dimesociety.org/playbooks/the-playbook/

[57] Bushmakin AG, Cappelleri JC. A Practical Approach to Quantitative Validation of Patient-Reported Outcomes. Wiley; 2022

[58] de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine. Cambridge University Press; 2011

[59] Butler J, Spertus JA, Bamber L, Khan MS, Roessig L, Vlajnic V, et al. Defining changes in physical limitation from the patient perspective: Insights from the VITALITY-HFpEF randomized trial. European Journal of Heart Failure. 2022;**24**(5):843-850

[60] US FDA. Incorporating Clinical Outcome Assessments into Endpoints for Regulatory Decision-Making [Internet]. U.S. Food and Drug Administration; 2019. Available from: https://www.fda.

gov/media/132505/download [Accessed: September 8, 2024]

[61] Di J, Bai JF, Isik K, Camerlingo N, Demanuele C. Deployment and application of multi-modal sensors in clinical trials. Biopharmaceutical Representative. 2022;**29**(1):6-10

[62] Yuda E, Yoshida Y, Sasanabe R, Tanaka H, Shiomi T, Hayano J. Sleep stage classification by a combination of actigraphic and heart rate signals. Journal of Low Power Electronics and Applications. 2017;**7**(4):28

[63] Lock EF, Hoadley KA, Marron JS, Nobel AB. Joint and individual variation explained (JIVE) for integrated analysis of multiple data types. The Annals of Applied Statistics [Internet]. 2013;**7**(1):523-542. Available from: http://projecteuclid.org/euclid.aoas/1365527209

[64] Di J, Spira A, Bai J, Urbanek J, Leroux A, Wu M, et al. Joint and individual representation of domains of physical activity, sleep, and circadian rhythmicity. Statistics in Biosciences [Internet]. 2019;**11**(2):371-402. Available from: http://link.springer.com/10.1007/s12561-019-09236-4

[65] Kang SJ, Leroux A, Guo W, Dey D, Strippoli MPF, Di J, et al. Integrative modeling of accelerometry-derived sleep, physical activity, and circadian rhythm domains with current or remitted major depression. JAMA Psychiatry [Internet]. 2024;**81**(9):911-918. Available from: https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2819864

[66] Ramsay J, Silverman BW. Functional Data Analysis. Springer Science & Business Media; 2005

[67] Matabuena M, Karas M, Riazati S, Caplan N, Hayes PR. Estimating knee movement patterns of recreational runners across training sessions using multilevel functional regression models. American Statistical [Internet]. 2023;**77**(2):169-181. DOI: 10.1080/00031305.2022.2105950

[68] Morris JS. Functional regression. Annual Review of Statistics and Its Application [Internet]. 2015;**2**:321-359. Available from: https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-010814-020413

[69] Hall P, Hosseini-Nasab M. On properties of functional principal components analysis. Journal of the Royal Statistical Society, Series B (Methodology) [Internet]. 2006;**68**(1):109-126. DOI: 10.1111/j.1467-9868.2005.00535.x

[70] Goldsmith J, Bobb J, Crainiceanu CM, Caffo B, Reich D. Penalized functional regression. Journal of Computational and Graphical Statistics [Internet]. 2011;**20**(4):830-851. DOI: 10.1198/jcgs.2010.10007

[71] Goldsmith J, Wand MP, Crainiceanu C. Functional regression via variational Bayes. Electronic Journal of Statistics [Internet]. 2011;**5**:572-602. Available from: https://projecteuclid.org/journals/electronic-journal-of-statistics/volume-5/issue-none/Functional-regression-via-variational-Bayes/10.1214/11-EJS619.full

[72] Smirnova E, Ciu E, Tabacu L, Leroux A. Scalar-on-function regression: Estimation and inference under complex survey designs. Statistics in Medicine [Internet]. 2024;**43**:4559-4574. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.10194

[73] Cui E, Crainiceanu CM, Leroux A. Additive functional cox model. Journal

of Computational and Graphical Statistics [Internet]. 2021;**30**(3):780-793. DOI: 10.1080/10618600.2020.1853550

[74] Cui E, Leroux A, Smirnova E, Crainiceanu CM. Fast univariate inference for longitudinal functional models. Journal of Computational and Graphical Statistics [Internet]. 2022;**31**(1):219-230. DOI: 10.1080/10618600.2021.1950006

[75] Cui E, Li R, Crainiceanu CM, Xiao L. Fast multilevel functional principal component analysis. Journal of Computational and Graphical Statistics [Internet]. 2023;**32**(2):366-377. DOI: 10.1080/10618600.2022.2115500

[76] Li R, Xiao L, Smirnova E, Cui E, Leroux A, Crainiceanu CM. Fixed-effects inference and tests of correlation for longitudinal functional data. Stat Med [Internet]. 2022;**41**(17):3349-3364. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.9421

[77] Agarwala N, Zipunnikov V, Crainiceanu C, Leroux A. Quantifying the time-varying association between objectively measured physical activity and mortality in US older adults over a 12-year follow-up period: The NHANES 2003-2006 study. BMJ Evidence-Based Medicine. 2024. DOI: 10.1136/bmjebm-2023-112303

[78] Schrack JA, Zipunnikov V, Goldsmith J, Bai J, Simonsick EM, Crainiceanu C, et al. Assessing the "physical cliff": Detailed quantification of age-related differences in daily patterns of physical activity. The Journals of Gerontology: Series A [Internet]. 2014;**69**(8):973-979. DOI: 10.1093/gerona/glt199

[79] Wrobel J, Muschelli J, Leroux A. Diurnal physical activity patterns across ages in a large UK based cohort: The UK biobank study. Sensors. 2021;**21**(4):1545

[80] Matabuena M, Petersen A, Vidal JC, Gude F. Glucodensities: A new representation of glucose profiles using distributional data analysis. Statistical Methods in Medical Research [Internet]. 2021;**30**(6):1445-1464. DOI: 10.1177/0962280221998064

[81] Goldsmith J, Crainiceanu CM, Caffo BS, Reich DS. Penalized functional regression analysis of white-matter tract profiles in multiple sclerosis. Neuroimage [Internet]. 2011;**57**(2):431-439. Available from: https://www.sciencedirect.com/science/article/pii/S1053811911004435

[82] Vu T, Seal S, Ghosh T, Ahmadian M, Wrobel J, Ghosh D. FunSpace: A functional and spatial analytic approach to cell imaging data using entropy measures. PLOS Computational Biology [Internet]. 2023;**19**(9):e1011490. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1011490

[83] Ballard ED, Greenstein D, Reiss PT, Crainiceanu CM, Cui E, Duncan WC, et al. Functional changes in sleep-related arousal after ketamine administration in individuals with treatment-resistant depression. Translational Psychiatry [Internet]. 2024;**14**(1):1-9. Available from: https://www.nature.com/articles/s41398-024-02956-2

[84] Shi JQ, Wang B, Will EJ, West RM. Mixed-effects Gaussian process functional regression models with application to dose–response curve prediction. Stat Med [Internet]. 2012;**31**(26):3165-3177. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.4502

[85] Goldsmith J, Scheipl F, Huang L, Wrobel J, Di C, Gellar J, et al. refund: Regression with Functional Data [Internet]. 2024. Available from: https://cran.r-project.org/web/packages/refund/index.html

[86] Crainiceanu CM, Goldsmith J, Leroux A, Cui E. Functional Data Analysis with R. Boca Raton: Chapman and Hall/CRC; 2024

[87] McCoy CE. Understanding the use of composite endpoints in clinical trials. Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health [Internet]. 2018;**19**(4):631-634. Available from: https://escholarship.org/uc/item/4693j9pv

[88] Sankoh AJ, Li H, D'Agostino RB Sr. Use of composite endpoints in clinical trials. Stat Med [Internet]. 2014;**33**(27):4709-4714. Available from: https://onlinelibrary.wiley.com/doi/10.1002/sim.6205

[89] Zarei S, Carr K, Reiley L, Diaz K, Guerra O, Altamirano PF, et al. A comprehensive review of amyotrophic lateral sclerosis. Surgical Neurology International [Internet]. 2015;**6**:171. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4653353/

[90] Johnson SA, Karas M, Burke KM, Straczkiewicz M, Scheier ZA, Clark AP, et al. Wearable device and smartphone data quantify ALS progression and may provide novel outcome measures. NPJ Digital Medicine [Internet]. 2023;**6**(1):1-10. Available from: https://www.nature.com/articles/s41746-023-00778-y

[91] Karas M, Olsen J, Straczkiewicz M, Johnson SA, Burke KM, Iwasaki S, et al. Tracking amyotrophic lateral sclerosis disease progression using passively collected smartphone sensor data. Annals of Clinical and Translational Neurology [Internet]. 2024;**11**(6):1380-1392. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/acn3.52050

[92] Straczkiewicz M, Karas M, Johnson SA, Burke KM, Scheier Z, Royse TB, et al. Upper limb movements as digital biomarkers in people with ALS. EBioMedicine [Internet]. 2024;**101**. Available from: https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(24)00071-9/fulltext

[93] van Eijk RPA, Bakers JNE, Bunte TM, de Fockert AJ, Eijkemans MJC, van den Berg LH. Accelerometry for remote monitoring of physical activity in amyotrophic lateral sclerosis: A longitudinal cohort study. Journal of Neurology. 2019;**266**(10):2387-2395

[94] Gupta AS, Patel S, Premasiri A, Vieira F. At-home wearables and machine learning sensitively capture disease progression in amyotrophic lateral sclerosis. Nature Communications [Internet]. 2023;**14**(1):5080. Available from: https://www.nature.com/articles/s41467-023-40917-3

[95] Neumann M, Kothare H, Ramanarayanan V. Combining multiple multimodal speech features into an interpretable index score for capturing disease progression in amyotrophic lateral sclerosis. Interspeech [Internet]. 2023;**2023**:2353-2357. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11246072/

[96] Di J, Demanuele C, Kettermann A, Karahanoglu FI, Cappelleri JC, Potter A, et al. Considerations to address missing data when deriving clinical trial endpoints from digital health technologies. Contemporary Clinical Trials [Internet]. 2022;**113**:106661. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1551714421003979

[97] Acharya J, Basu A. Deep neural network for respiratory sound classification in wearable devices enabled by patient specific model tuning. IEEE Transactions on Biomedical Circuits and Systems. 2020;**14**:535-544

[98] Hssayeni MD, Jimenez-Shahed J, Burack MA, Ghoraani B. Wearable sensors for estimation of parkinsonian tremor severity during free body movements. Sensors. 2019;**19**(19):4215

[99] Yang K, Ahn CR, Vuran MC, Aria SS. Semi-supervised near-miss fall detection for ironworkers with a wearable inertial measurement unit. Automation in Construction. 2016;**68**:194-202