Therapeutic Innovation & Regulatory Science
https://doi.org/10.1007/s43441-020-00187-3

DIA

ANALYTICAL REPORT

Determining Minimum Wear Time for Mobile Sensor Technology

Marie McCarthy, MSc, MBA[1] · Denise P. Bury, MPH, PhD[2] · Bill Byrom[3] · Cindy Geoghegan[4] · Susan Wong, PA[5]

Received: 20 December 2019 / Accepted: 9 June 2020

Abstract
Part 1 in the DIA Study Endpoint Community Working Group on Mobile Sensor Technology (MST) series addresses considerations that may be useful when determining the minimum wear time associated with mobile sensor use to ensure reliable estimation of the clinical endpoint under consideration. What constitutes a minimum valid data set is a dilemma facing those using MSTs in clinical studies. If this alignment does not occur, the integrity of the data collected and conclusions drawn from these data may be in incorrect. While study participants should consent to engage with MSTs as defined in a protocol, participant behavior or technology lapses may result in capturing incomplete data. Drawing from the literature, we review what constitutes a minimum data set, the risks associated with missing data, alignment with the clinical endpoint(s) and goals of a study, as well as managing patient burden.

Keywords Mobile health · Sensors · Wearable devices · Clinical trials · Minimum wear time · Valid data set · Digital health technologies

Introduction

This is the 1st of a series introduced by the DIA Study Endpoint Community Working Group on Mobile Sensor Technology tasked with investigating key considerations for the use of mobile sensor technology (MST) (non-drug delivery) in a clinical trial (CT) setting. This working group was convened among members of the Drug Information Association's (DIA) Study Endpoint Community including experts from biopharmaceutical companies, eClinical technology providers, clinical research organizations (CROs), academia, and non-profit organizations. The working group seeks to provide further recommendations and guidelines to facilitate the adoption of mobile sensor technology in clinical trials, allowing for the objective demonstration of treatment benefit in a real-world setting. This manuscript addresses considerations regarding the determination of minimum wear time for MSTs used to generate clinical endpoints.

The goal of each clinical study is to generate data to support a hypothesis. Endpoints should adequately measure the concept(s) of interest of the study, and endpoint data need to be reliable, robust, and sufficiently powered to be statistically significant. As with any other outcome measure, if sufficient data is not collected, the reliability of the endpoint estimate and the statistical power of the data is called into question. The majority of MSTs are used to capture data remotely, and while the value of this continuous data capture methodology has been well established, there is no standard approach to determine the amount of time a study participant should wear or interact with the technology to ensure robust and reliable endpoint estimates. The impact and significance of wear compliance has not been quantified; that is, the percentage of that pre-defined portion of the time that study participant wears or interacts with the technology. A compounding factor in determining standards in the use of MSTs is the significance of the potential data loss that is dependent on the use of a specific MST in a given study, study population and therapeutic area, endpoint derived, and role of the MST in the study. In addition, should MSTs be used for longer than required, there could be significant cost implications in terms of the number of devices required,

✉ Marie McCarthy
Marie.mccarthy@iconplc.com

1 ICON PLC, South County Business Park, Leopardstown, Dublin 18, Ireland

2 Sanofi-Aventis Groupe (Sanofi), Cambridge, USA

3 Signant Health, London, England, UK

4 Patient and Partners LLC, Madison, CT, USA

5 Transparency Life Sciences, Inc., Philadelphia, USA

Published online: 25 June 2020

additional burden on study participants, and increased data complexity and noise.

What Constitutes a Minimum Valid Data Set for Mobile Sensor Technologies?

What constitutes a valid data set with respect to the use of MSTs is not well defined. While the goal is to optimize wear/usage time, to accurately and comprehensively capture the target endpoint, there are a number of alternative approaches. A valid data set, on a participant level, might be defined as the amount of data required to enable the estimation of the health outcome(s) of interest to a degree of accuracy deemed appropriate. Equally, a valid data set might be considered within a specific assessment interval; for example, how many hours of wear time per day would be sufficient to obtain a reliable estimate of the health outcome(s) of interest.

In general, it can be difficult to determine absolute values from the literature, although for certain types of MSTs, such as accelerometers to measure activity or sleep, some scientific evidence does exist. A valid data set in terms of wear time cannot be considered simply by a device type, as the target clinical endpoint derived from the MST data is also related to the amount of wear time needed to generate a valid data set. For example, in the literature, to assess overall daily activity, such as total steps per day, typically, researchers assume that a minimum of 80% of data need to be available (non-missing) in order to have a valid data set [1]. Despite this recommendation being widely used, it is open to interpretation. For example, 80% non-missing data could refer to 5.6 days out of a requested 7 days of data, or, alternatively, could correspond to at least 19.2 h for each day within a 7-day period. This brings us to the distinction between defining a valid day of wear and establishing a minimum number of valid days required. A valid day wear interval ensures that wear time is sufficient to determine a reliable estimate of the clinical endpoint for that day. The minimum number of valid days ensures that, allowing for day-to-day variability, we are able to obtain a reliable daily mean estimate of the endpoint.

Minimum wear time will vary based on the endpoint to be measured, For example, if using an activity monitor to measure total steps per day, between 6 and 16 h wear time during a day has been reported to provide acceptable estimates of daily steps during the active-awake period [2]. Chen et al. explored the definition of a "usable day" as being either 6, 8, or 10 h of estimated wear time and compared the impact of the definition on the sensitivity of estimation of time spent in Moderate and Vigorous Physical activity (MVPA) in a study of 1685 participants. The results of this study suggested that using a valid day definition as low as 6-h wear time had little impact on the estimate of mean MVPA minutes and was similar to the normalized wear time of 12 h [3]. However, measuring a different endpoint may have a significant impact on the definition of valid days required for robust estimation. Using the same device to measure real-world walking speed, for example, it is possible that this can be estimated robustly from just a small number of bouts of purposeful walking and therefore may not require the same amount for daily wear time [4].

In order to determine the number of valid days that enable a robust definition of overall activity, the lifestyle, age, and work patterns of study participants need to be considered, as will an understanding of the variability in activity between days. The average adult study participant that typically works Monday to Friday may be more active during the days on weekends compared to weekdays [5], whereas retired or home-bound patients may have similar activity levels on all days of the week. As a result, some researchers report that only 3 days of accelerometer data are required to accurately predict physical activity in older adults, whereas 5 consecutive days or 6 randomly selected days of day is required in adults (mean age 38 ± 10 years) to account for the potentially significant differences in activity patterns [6].

Similar approaches are required when using MSTs to estimate other physiological measures—for example, measuring FEV1 or oximetry levels in patients living with COPD. In this case, understanding valid days and the minimum number of days needed remains important to ensure that endpoint estimates can be seen as reliable, and that additional variability does not adversely impact the sensitivity of the endpoint to detect change when change exists.

Approaches to the Standardization of a Valid Data Set

When determining minimum wear time for a study, the data to support the definition of wear time requirements can be obtained from pilot studies, scientific literature or other large data sets [7], and through engaging relevant patient perspectives. Considerations should include the endpoint under-consideration. For example, actigraphy-derived sleep parameter estimation wear time should align with pre-existing established clinical practice as outlined in the International Classification of Sleep Disorders (ICSD), the primary reference tomb for Sleep researchers. The ICSD outlines the appropriate use of actigraphy dependent on the specific population and sleep disorder under investigation [8]. The clinical use of physical activity monitors is less common, and inferences can be drawn from the research community and large community-based population studies such as National Health and Nutrition Examination Survey (NHANES), UK Biobank, or Canadian Health Measures

Survey. In the absence of clinical guidance, the periods during a specific day that are required to obtain a reasonable estimate of activity over the day should be considered, and as described above, this may be dependent on the specific population under consideration. Typically, days should be standardized to a common active time interval period, such as 16 h active wake time, and intervals above or below this value are adjusted proportionally. Evaluating data sources to enable the sensitivity of the endpoint estimate to wear time will be valuable in determining the optimal minimum wear interval to define a valid day.

There is significant value in the use of pilot studies with representative patient populations. These studies facilitate an understanding of the impact and wear compliance allowing the study team to benchmark and set minimum wear tolerances and to test assumptions to ensure the most suitable limits for the study, population, outcome assessment, and protocol are established. Minimum wear time prior to trial initiation may be a requirement in some cases, specifically, to determine the amount of time that patients need to wear the sensor prior to baseline to measure change in endpoint.

When data is not readily available from other sources, findings reported in the literature can be applied to the use of other sensors within a specific device family. For this approach to have value, one needs to standardize to a common denominator. This approach was used by Byrom et al. [9] in their review of activity monitoring in COPD studies in which they determined from the published methodologies and results that the denominator should be 16 h if looking at total daily activity measures. Hermann et al. used data sets from the NHANES to explore the impact of wear time on activity estimates. This study explored data from 4000 individuals and evaluated the impact of reducing wear time on the robustness of activity estimates. This study concluded that using 12 h or less wear time data significantly reduced the estimates of time spent in activity and sedentary behavior [10]. However, this approach is specific for this disease area and endpoint, and alternative approaches are required when investigating other patient populations or different derived clinical endpoints.

Missing Data

Defined by Little et al. as "values that are not readily available that would be meaningful for analysis if they were observed" [11], missing data poses one of the greatest threats to data validity and generalizability of results regardless of whether the data comes from MSTs or other sources. Missing data has significant implications and can lead to incorrect inferences, reduce statistical power, may introduce biases, and may impact the representativeness of the study sample [12]. This could have the impact or reaching an inappropriate study conclusion or failing to flag drug safety issue [11]. While there is lack of consensus within the industry as to whether risk determination is a data management or a function of biostatistics, what is clear is that missingness poses a risk to the ability to draw and support conclusions relating to study results associated with the intended endpoint.

Within the context of MSTs, this issue is even more acute. The majority of wearables and sensors are used remotely. The value of this continuous data has been well established, but the impact on a clinical study of incomplete data, where study participants interact or wear the devices for a portion of the required time, is less well established. When data are collected in unsupervised settings, it can be difficult to understand the reason for missingness. An understanding as to the circumstance of the missing data can be important in the ability to determine the appropriateness of possible rules for dealing with missing data. Lack of knowledge of whether missing wear time is missing at random, or not at random, makes it difficult to develop robust approaches to interpolating/estimating missing values. Missing data at random (MAR) could be due to the malfunction of the technology or a patient simply forgetting to wear the device for a period of the day. Missing data not at random (MNAR) could be a result of the patient's illness and feeling unable to wear or use a device or sensor or removing the device to perform a specific activity (e.g., swimming). This "not at random" incidence may, in this case, represent a period of inactivity which would be important to measure [13]. Strategies on handling missing and incomplete data should be selected based on the reason(s) for missing and incomplete data and the patterns of missing data. For data that are classified as MAR, imputation methods may be appropriate, as it may be possible to assume that the data from the study days with incomplete data will have the same distribution as the complete days [14]. However, when data are MNAR, a pattern mixture modeling approach is required. When data are assumed to be MNAR, those patients who complete the assessments may be the healthiest, while those who have progressively become sicker may have dropped out of study. The ability to determine MAR or MNAR can impact bias if imputation is used where these techniques have been shown to be less robust in eliminating bias in the overall summary measure of activity [1].

Once the minimum wear time sufficient to clinical endpoint has been determined, it is realistic to suppose that established strategies for handling missing data could be applied to MST data. In particular, depending on the proportion of missing data, multiple imputations could play a role, and recent improvements in multiple imputation and statistical computing strategies are assisting statisticians with techniques to address the issue of missing data [15]. Despite this, there are few examples of the use of imputation methods to

account for missing data and researchers typically discard data that does not meet a valid day definition [9].

An alternative approach could be the use of Quality Tolerance Limits (QTL), a risk-based monitoring approach that is gaining momentum in the industry. The ultimate aim of QTL is to identify systemic issues at trial level that have the potential to impact subject safety or data integrity. In accordance with ICH E6 (R2) [16], each sponsor determines the acceptable level of risk and defines the limits or tolerance for each specific identified parameter. These tolerance limits are pre-specified by study sponsors and are designed to minimize risks to subject safety and ensure the reliability of the study results. For example, in a study where MSTs were used, a sponsor would predefine compliance level above which a protocol violation would be triggered. Early identification of issues can trigger an investigation and/or mitigation strategies to improve QTLs. The establishment of QTL's for MST endpoints has the potential to identify significant issues with the data and it is hoped that these QTLs have the potential to become more meaningful as the use of MSTs in clinical trials matures [17].

Patient Factors

Once critical clinical research requirements are established, it is also essential to engage and collect perspectives and insights from relevant patient populations. This has the benefit of ensuring that specified wear time and device usability are acceptable and that potential participant concerns are addressed during planning. Incorporating these perspectives from the relevant patient populations is essential when determining wear time and device usability and acceptability for the specified wear time. Patients should understand how expectations for wear time throughout their study participation may differ from their experiences with consumer tools and wearables. There are other issues beyond the form factor of an MST that have the potential to impact wear, including the availability for technical support should participants perceive a device is malfunctioning as well as perceptions related to value of being able to access personal data in real time. Perceived risks to privacy and security need to be addressed. However, despite these concerns, a recent survey of 200 potential clinical trial participants investigating preferences and perspectives of patients in deciding whether to participate in a mobile clinical trial [18] reported that respondents overwhelmingly chose a mobile trial scenario vignette over a traditional trial approach, citing convenience, time, and access as benefits. Respondents also said they would wear a device or monitor for a year or more—as long as the trial lasted.

Collecting more data than is necessary/ irrelevant to the primary endpoint may put undue burden on participants. Engaging patients early in planning clinical trials using MSTs will help ensure that collecting data remotely and continuously provides benefit and value to participants while balancing any added burdens related to wearing and maintaining devices and learning to live with a new technology. Seeking input from representative patients may also predict situations under which data capture may be inconvenient and less relevant to patient outcomes, thereby potentially reducing instances of missing data. Additionally, patients may be open to using MSTs for longer than might be expected if they are included in planning trials and understand why data are essential to potential better outcomes.

Discussion and Conclusion

When considering using MSTs in a study, there are three factors that need to be deliberated; defining what is meant as a valid data set, optimizing data collection from study participants in terms of complete data sets as defined by the protocol, and utilizing computational strategies for managing missing data.

Guidance needs to be provided during the study design period to determine the definition of valid data, rules for handling missing data, and when/if data should be judged unreliable and excluded. Determining minimum wear time for robust estimates of clinical endpoints is critical. Data from the literature are at best broad and few studies to date provide evidence to link the wear time to specific clinical endpoints and patient populations and it is likely that until a sufficient body of research has been conducted across a number of therapeutic areas with different patient cohorts, the definition of a valid data set will be broadly based with generalities such as 80% of the intended wear time.

In concert with the need to define a valid data set and the proportion of that data that is missing, maximizing the wear compliance of study participants is essential. This can be achieved through engaging with study participants to gain their perspective, addressing their preferences and acceptability of MSTs to alleviate non-adherence, and to provide technical support and training to maximize engagement with the MST. These initiatives can include pilots to help benchmark wear tolerances levels, real-time compliance programs, active outreach, and post-study questionnaires. From a patient's perspective, it is critical to balance additional burden and risk (privacy, wear time and device failure) with returning value to participants.

Mobile sensor technologies are increasingly being used in clinical studies, but there remain unknowns and is it likely that approaches to the management of missing data and definition of a valid data set will evolve and become more precise and defined as these tools become more widely adopted. Once statisticians are able to distinguish between

data missing at random and not missing at random, well-established techniques such as and multiple imputation and advanced computational programs can be used to address the issue of missing data.

While some questions are still yet unanswered, such as how much data is required for that data to be considered a valid data set and included in the statistical analysis and how rigid should the statistical analysis plan (SAP) be when defining a valid data set, there is still enough known and sufficient insights that can be derived using MSTs to encourage greater use of MSTs in clinical trials. Existing approaches already utilized by bio-statisticians have a role in determining the impact of the missing data such as sensitivity analysis that could be conducted to assess the robustness of findings to plausible alternative assumptions about the missing data [11]. In the meantime, the industry has adopted broad strategies with respect to wear time with an expectation that more specific guidelines will evolve as industry more accurately defines more precise values for what constitutes a valid data set for each specific use case and indication.

Funding
There were no funding sources. No financial support was received.

Compliance with Ethical Standards

Conflict of interest
There were no conflicts of interest.

References

1. Catellier DJ, Hannan PJ, Murray DM, et al. Imputation of mission data when measuring physical activity by accelerometry. *Med Sci Sports Exerc*. 2005;37(11):S555–62.
2. Herrmann SD, Barreira TV, Kang M, Ainsworth BE. How many hours are enough? Accelerometer wear time may provide bias in daily activity estimates. *J Phys Activity Health*. 2013;10:742–9.
3. Chen C, Jerome GJ, Laferriere D, Young DR, Vollmer WM. Procedures used to standardize data collected by RT3 triaxial accelerometers in a large-scale weight-loss trial. *J Phys Act Health*. 2009;6(3):354–9.
4. Orendurff MS, Schoen JA, Bernatz GC, Segal AD. How humans walk, bout duration, steps per bout, rest duration. *J Rehabil Res Dev*. 2007;45(7):1077–89.
5. Drenowatz C, Gribben N, Wirth MD, Hand GA, Shook RP, Burgess S, Blair SN. The association of physical activity during weekdays and weekend with body composition in young adults. *J Obes*. 2016;2016:8236439.
6. Hart TL, Swartz AM, Cashin SE, Strath SJ. How many days of monitoring predict physical activity and sedentary behaviour in older adults? *Int J Behav Nutr Phys Act*. 2011;8:62.
7. Doherty A, Jackson D, Hammerla N, et al. Large scale population assessment of physical activity using wrist worn accelerometers: the UK Biobank study. *PLoS ONE*. 2017;12(2):e0169649.
8. American Academy of Sleep Medicine. 2014. International classification of sleep disorders.
9. Byrom B, Rowe DA. Measuring free-living physical activity in COPD patients: deriving methodology standards for clinical trials through a review of research studies. *Contemp Clin Trials*. 2016;47:172–84.
10. Herrmann SD, Barreira TV, Kang M, Ainsworth BE. Impact of accelerometer wear time on physical activity data: a NHANES semi-simulation data approach. *Br J Sports Med*. 2014;48(3):278–82.
11. Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med*. 2012;367:1355–60.
12. Kang H. The prevention and handling of the missing data. *Korean J Anesthesiol*. 2013;64(5):402–6.
13. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: from design to analysis. *Yale J Biol Med*. 2013;86:343–58.
14. Rubin DB. Inference and missing data. *Biometrika*. 1976;63:581–92.
15. Belin TR. Missing data: what a little can do and what researchers can do in response. *Am J Ophthalmol*. 2009;148(6):820–2.
16. E6(R2) Good Clinical Practice: Integrated Addendum to ICH E6 (R1) Guidance for Industry. US Department of Health and Human Services, Food and Drug Administration. March 2018. www.fda.gov/media/93884/download. Accessed 13 May 2020.
17. TransCelerate Biopharma, Inc. Risk-based quality management: quality tolerance limits and risk reporting. 2017.
18. Perry B, Geoghegan C, Lin L, et al. Patient preferences for using mobile technologies in clinical trials. *Contemp Clin Trials Commun*. 2019;15:100399.