PharmaSUG 2023 - Paper SD-300

Standardized data handling framework for wearables

Madhu Annamalai - Technical Lead, Statistical Programming, Algorics
Umayal Annamalai - Assistant Manager, Statistical Programming, Algorics

ABSTRACT

Wearable devices have revolutionized the way clinical research is conducted in recent times, thereby bringing technology closer to the patients. While wearables can help continuous and near-real-time monitoring of patient data, analysis and reporting continue to be challenging as the quantum of data continues to grow. Problem and solutions framework for handling data from wearables to improve analysis and reporting is continuously changing and challenging. The key takeaways will include data integration strategies, data mapping based on CDISC standards, creating wearable device data-specific analysis models, and the generation of custom reports based on the study requirements.

INTRODUCTION

Direct data capture (DDC) methods using digital data sources have been used widely in the current landscape of clinical trials. Of those digital data sources, wearable device data help enhance patient surveillance on a near-time basis. Wearables are widely used to track patients' health, sleep, and activity data. Parameters such as heart rate, respiratory rate, blood oxygen saturation level, and body posture can be collected seamlessly. For instance, an accelerometer embedded in a wristband is an example of a sensor passively collecting data about a person's physical activity and movement. This is a big step forward compared to the traditional means of health-related data collection within Clinical trials. Table 1 shows several types of wearable data (with distinct brands) and the type of data being collected.

Table 1: Examples of Wearable Devices

Device types	Wearable Devices	Data Collected
• Wrist- Mounted • Ankle -Mounted • Waist -Mounted • Ear Mounted • Skin-Patch	• Fitbit Charge HR/Flex • Garmin Vivo fit 2/3 • ACTi graph GT3X or wGT3X-BT • Mini Mitter Phillips ACTi watch 64 • Jawbone UP • Mi Band • Apple watch – Samsung Gear • Active Tracer AC-301 • Fitbit Zip [Clip-on] • Trium ACTi belt • Active Style Pro HJA-350IT • PAM Model AM101 • e-AR ear-worn activity • STO2 Device	• TDS (Total Daily Steps) PAI (Physical Activity Intensity) • CA (Cumulative Acceleration) • Total steps ambulated. • TDDT (total distance walked) • Gait velocity/speed • Step Length • DCE (Daily Caloric Expenditure) • PAM Score [Personal Activity Monitor] • Tissue Oxygen Saturation [STO2]

Challenges in Data standardization and Reporting in Wearable Data

While using wearable has many advantages, the huge quantum of data poses challenges in data handling and processing given that lack of standardization. Below are the key issues that exist sporadically in the current scenario.

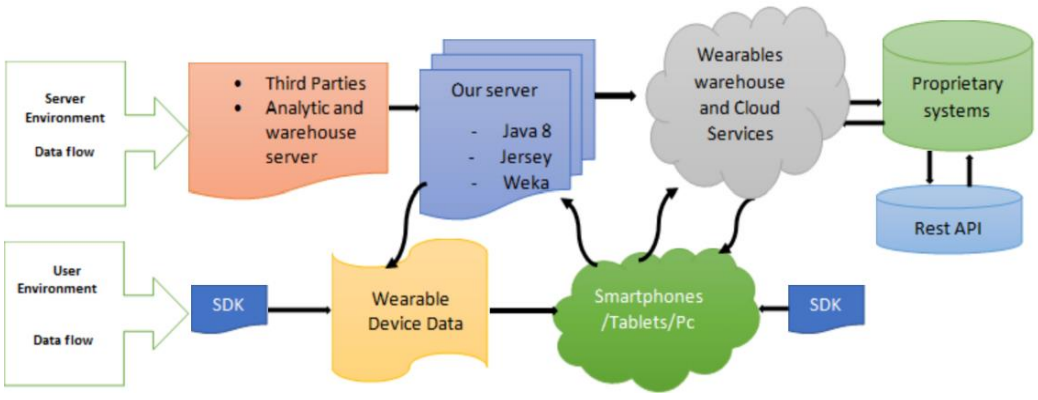
- Cost, time, and infrastructure limitations involved in the adaptability of wearable devices.
- Cybersecurity vulnerabilities and device malfunction
- Literature review for scoring algorithms and data pre-processing is limited.
- Daylight savings transition – Data handling becomes challenging from processing to reporting.
- Data mapping in SDTM is challenging as we have an enormous number of tests with trivial differences.
- ADaM Derived parameters are quite complex.
- Data derivations in ADaMs, namely – valid day, valid week, valid baseline data, and analysis flags for data summary tables, etc are challenging.
- Visit windowing for the Activity/sleep data, which are collected daily. Few differences in visit windowing conditions will tweak the valid data conditions.
- Compliance calculations are simple, but Compliance is tested in baseline and has impacts on valid data derivations.
- Planning of analysis datasets needs detailed attention to facilitate TFL programming.

To tackle the above challenges, we are proposing different perspectives starting from data acquisition & processing, data standardization and reporting of wearable data in the sections below of this paper.

Data Acquisition & Processing:

A clear data acquisition strategy, including a data flow map and integration, helps consolidate the data collected from wearable devices. Patient data collected through wearable devices will need to be integrated as e-health records with device serial numbers and saved in an intermediate data system (or like a Data Lake). After indigestion, the data can be transformed as needed for further processing. After data ingestion, data cleaning & preprocessing, such as the validity of the data, device compliance, and continuity of the data for meaningful analysis, must be performed.

Figure 1: Data Acquisition



Additional checks may include (1) trimming of non-data from the beginning and end of a file, which keeps or marks the section of data used for analysis, (2) identifying and marking 'off-wrist' segments (i.e., non-wear periods) to exclude these intervals from analyses (3) considerations around data where daylight savings are happening (4) considerations while the patient is working in shifts specifically where the sleep data could be different from the regular sleep pattern. Any scoring methodology to be used for derivations needs to be verified and validated.

Data Standardization:

Once the data processing is completed, the data is ready to be transformed to CDISC SDTM datasets for further analysis. Clear SDTM mapping specification will help to handle the number of --TEST in the datasets. Once the SDTM transformation is completed, analysis variables will be prepared, which are ready for analysis based on the ADAM data model (custom BDS domain). Examples of ADAM variables, including standard and derived variables, are given below:

- Identify the valid activity data for the day (>=xx hours of wear time)
- Identify the valid sleep data for the day (>=xx hours of awake wear time)
- Identify the valid data for the week (xx no of valid days of data within the week)
- Summarize the mean of the valid data within the week as derived data (Conversion of daily data to weekly data (consolidated data based on study reporting requirements) (baseline data, post-baseline data will have to go through these steps)
- Derivations on baseline change over time.
- The derivation is needed for record level, subject level, digital data analysis flags, and post-baseline records flags as needed for the TFLs.

Additionally, compliance dataset based on ADaM model should be generated to summarize the compliance of the wear conditions based on the study requirements.

Data reporting

Following data standardization, the data reporting elements will need to be determined. These data elements may include baseline data table, change from baseline table, mean, and line plots, plots for correlation matrix between the activity and sleep parameters, device wear compliance analysis, model-based statistical analysis with/without co-variates/stratum and subgroup analysis upon the endpoints. But this is not limited to the above list, a lot more extensive research could be required based on study requirements as the wearable devices-based endpoints are recognized as exploratory endpoints. Hence, all the reporting requirements should be handled in the analysis datasets effectively to generate the required reports seamlessly.

CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH

Wearable devices have numerous advantages in the current era of clinical digitalization world. However, meticulous planning and a detailed strategy are required for the design, collection, standardization, and reporting of wearables. During the planning stage, the objective of using wearables in the context of the study endpoints will need to be defined. Further, a robust technology platform would help the seamless ingestion and integration of digital data into a centralized location. A centralized review of the data and endpoints would enable faster clinical decision-making. Finally, creating a standard framework for analyzing and reporting wearable devices data upon digital endpoints based on CDISCs structure (SDTM/SDTM-like/ADaM) would support the regulatory submission. Consequently, a constant evolution of this topic is much needed for fine-tuning the approach and effective usage of wearable data for prospective clinical trials.

REFERENCES

- Bradford, Jennifer. 2019. " The promise of wearable technology in clinical trials: a proof of concept to deliver it? " *Proceedings of the Phuse 2019*, RW02.
- Fekedulegn, D., Andrew, M. E., Shi, M., Violanti, J. M., Knox, S., & Innes, K. E. 2020. "Actigraphy-Based Assessment of Sleep Parameters." *Annals of Work Exposures and Health,* 64(4), 350-367. Available at https://doi.org/10.1093/annweh/wxaa007.

ACKNOWLEDGMENTS

We would like to thank Mukesh Babu K, Head of Clinical & Data Operations for support and guidance.

RECOMMENDED READING

- *Wearable Technologies by Nicola Carbonaro; Alessandro Tognetti, January 2019*
- *Data Mining and Predictive analysis second edition by Colleen McCue,2015*

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Umayal Annamalai
Assistant Manager, Statistical programming,
Algorics
Email: Umayal.Annamalai@algorics.com

Madhu Annamalai
Technical Lead, Statistical programming,
Algorics
Email: Madhu.Annamalai@algorics.com