DEPLOYMENT AND APPLICATION OF MULTI-MODAL SENSORS IN CLINICAL TRIALS

Junrui Di (Pfizer), Jiawei Bai (Pfizer), F. Isik Karahanoglu (Pfizer), Nunzio Camerlingo (Pfizer), Charmaine Demanuele (Pfizer)

1. Introduction

Over the last decade, Digital Health Technologies (DHTs) have proven to be effective in measuring human activity and physiology, and been integrated into numerous clinical trials (Digital Medicine Society (DiME), 2021). DHTs have evolved rapidly especially in recent years, with astonishing strides toward miniaturization and life-cycle extension. It has been reported previously (Karas *et al.*, 2019) that small sensors such as accelerometers could provide valuable information about the study participants' health condition. While more of these sensors (e.g. heart rate monitor, thermometer, pulse oximeter (SpO2), continuous glucose monitoring) were being packaged together in modern wearable devices, one of the most common questions to ask is, whether and how we can better understand human health using the data from all these sensors combined. This question arises naturally from the fact that each sensor often provides location-specific information, that is, limited to one dimension of human activity or physiology. For example, accelerometers measure the magnitude of physical movement of the body, while heart rate monitors may indicate the level of exertion. However, neither of these two sensors could reliably provide blood oxygen level, which would need an SpO2 sensor. For a study focusing on heart failure, each of the three domains (activity, exertion, and blood oxygen level) have its own clinical indication and missing any of them may fail to provide a holistic picture of disease progression. Therefore, combining multi-modal sensor data may enable more comprehensive phenotyping, better symptom characterization and more accurate assessment of changes in health status over time.

A few notable studies have explored this path. Merikangas *et al*. (2019) included both an accelerometer and an ecological momentary assessment (EMA) component in their study to examine the associations among motor activity, energy, mood, and sleep. The joint modeling of a) the motor activity and sleep measurement derived from accelerometry data and b) the mood and energy level assessed through the EMA devices offered the authors opportunities to gain insights into how these different domains interact with each other and potentially what the therapeutic target is for patients with bipolar disorder. A similar example is the Apple Women's Health Study (Mahalingaiah *et al*., 2021), which aims to investigate the relationship among women's menstrual cycles, health and behavior, through a "mobile-application-based longitudinal cohort study" that has both a sensor and a survey component. Their analysis aims to combine both the (monthly) survey data with longitudinally measured smartphone/watch data, so that a better understanding might be reached of how the menstrual cycle relates to exercise, sleep, environment, behavioral and other physiological processes. Besides typical observational studies, Quer *et al*. (2021) showcased that multi-modal sensor data including heart rate, sleep and activity coupled with self-reported symptoms could significantly distinguish between symptomatic individuals with and without a diagnosis of COVID-19

These studies all highlighted the fact that multi-modal sensors were beneficial because each sensor contributed distinct aspect of information to the statistical model. However, there are considerations researchers should be aware of, before conducting studies with multi-modal sensors. In the remainder of this article, we will first discuss typical analytical challenges, and then elaborate on the requirements of deploying multi-modal sensors in clinical studies.

2. Analytical Challenges of Multi-Modal Sensor Data and Emerging Techniques

Modern DHTs collect data passively, continuously, and frequently, leading to rich streams of time series data with high dimensionality, complex data structure, and potentially noisy signals. With features derived from multi-modal sensors, one can directly combine those features in linear or nonlinear fashion using statistical and machine learning models. For example, as discussed previously, by combining motion-related features acquired from actigraphy, and heart rate/heart rate variability features acquired from wearable electrocardiogram (ECG), the accuracy for sleep prediction and sleep stage classification can be potentially increased when compared to using only one of these modalities (Aktaruzzaman *et al*., 2017; Yuda *et al*., 2017). However, when multi-modal sensors are deployed simultaneously, some new challenges arise due to the continuous nature of the measurements, and the interrelation between different modalities. It becomes crucial to fully utilize the rich data and identify the homogenous underlying signals (such as disease progression or treatment effects) from multiple modalities while accounting for the possible heterogeneity across modalities.

2.1 Fully Utilize the Temporal Aspect of Sensor Data

Before fusing data collected by multiple sensor modalities, a key issue to consider is to leverage the continuous time series signals from each of the sensors. It is still common practice to derive features that quantify certain physiological or behavioral characteristics (Di *et al*., 2019). For example, total activity counts have been used to represent overall daily activity intensity in many studies using accelerometers (Varma *et al*., 2017). Similarly, time-in-ranges indices are commonly employed in studies involving continuous glucose monitoring (CGM) sensors to quantify the quality of glucose control (Battelino *et al*., 2019) However, these features are summary measures and do not reveal the temporal variations within a day.

In circadian rhythm research, cosinor and extended cosinor models have been utilized to parametrically estimate the daily diurnal trend as a cosinor (or transformed cosinor) curve to represent the amplitude and phase of time series data (i.e. time to reach the peak) (Marler *et al*., 2006; Cornelissen, 2014). Time series data collected by sensors can be considered as a function of time. More recently, functional data analysis (Georgiev *et al*., 1998), which was developed to study the smooth functional behaviors of curves over a continuum, has been widely use to nonparametrically estimate the temporal characteristics of physiological trends or diurnal patterns (Goldsmith *et al*., 2016). By assuming the underlying functional smoothness, functional data analysis approaches such as functional regression (function-on-scalar or scalar-on-function) and functional principal component analyses can identify treatment effects within a specific time window in a day, or to detect a shift of phase across different cohorts. For example, Spira *et al*. recently discovered significant differences in activity levels between participants with and without β-amyloid (Aβ) antibody only within specific time windows during a day, by using function-on-scalar regression (2021).

Other than emerging statistical methodologies that aims to reveal temporal trends, modern deep learning architectures such as Recurrent Neural Network (RNN) can ingest time-sequential data collected by wearables to solve for prediction problems (Nweke *et al*., 2018), such as human activity recognition (Chen *et al*., 2021). Specifically, RNN models using Long Short-Term Memory (LSTM) with different memory units have been widely used to model data collected by wearable devices (Rabby *et al*., 2021; Uddin and Soylu, 2021).

The prediction ahead of time of glucose concentration levels can be reliably achieved by exploiting their recent history, monitored by (minimally invasive or non-invasive) CGM sensors, in combination with data-driven algorithms. Simple data-driven strategies, using polynomial or linear autoregressive models (Eren-Oruklu *et al*., 2009), as well as more sophisticated methods, such as Kalman filters (Facchinetti *et al*., 2011) or neural networks (Rabby *et al*., 2021), have proven effective in the short-term prediction of future glucose levels (Prendin *et al*., 2021).

With the amount of available data collected by wearables rapidly growing, these deep learning

approaches and model-based techniques will become mainstream and standard approaches to deal with real-world measurements.

2.2 Separate the Joint Effects and Individual Modal Specific Effects

In clinical trials, fusion of multi-modal sensor data can be used to identify overall treatment effects by aggregating information from different physiological/behavioral domains. To reveal such effects, sometimes it is necessary to separate the joint effects that are homogenous across different modalities from the modal-specific effects.

In 2013, Lock *et al*., developed the Joint and Individual Variation Explained (JIVE) and used it to study the association between gene expression and miRNA data collected from the same samples (2013). As a data fusion technique and an extension to principal component analysis, JIVE decomposes 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. Di *et al*. applied it to integrate accelerometry-derived features quantifying three physiological domains of activity, sleep, and circadian rhythm quantify and separate between- and within-domain variation (Di *et al*., 2019). The same concept can be directly implemented to study features derived from multi-modal sensors. With recent generalizations and extensions such as to account for heterogeneous data types (continuous/binary/count) (Li and Gaynanova, 2017) and partially shared information between modals (Gaynanova and Li, 2017), JIVE shows the promise to fuse multi-modal sensor data.

JIVE provides a framework to properly quantify the interrelation and codependency across multiple data modalities. Conceptually, the interrelation and codependency can be considered as an outcome measurement by itself. With longitudinal clinical trials with multi-modal sensors, the change of such interrelation can be traced and analyzed to provide meaningful clinical interpretation.

3. Practical Considerations to Incorporate multi-modal sensors into clinical trials

FDA recently released the draft guidance "Digital Health Technologies for Remote Data Acquisition in Clinical Investigations" which provided recommendations on the use of DHTs in clinical investigations, such as considerations for device selection, endpoints validation and verification, and statistical analysis. (US FDA, 2021). Di *et al*. provided operational suggestions to deploy DHTs in clinical studies to minimize the impact of missing data (Di *et al*., 2022). Incorporating multi-modal sensors should in principle follow these suggestions, such as to configure the devices appropriately, to determine the optimal placement location of the device, and to collect additional contextual information, when possible.

For clinical studies where patients wear one or multiple devices for a long period of time, it is crucial to incorporate the patients' perspective to increase their adherence. At the design phase of the studies, focus group of patients can be used to capture their voice to understand the preferrable form factor of the device(s) and the outcome measures that is the most meaningful to their daily life and health conditions. One question that can be considered is that to obtain a holistic view of multiple physiological/cognitive/behavior/environmental domains, should we identify a single device with multiple embedded sensors instead of providing multiple devices? To reduce risks to patients, as suggested by FDA in the draft guidance (US FDA, 2021), it is important to have a comprehensive informed consent of human subjects that details what data will be acquired from the multi-modal sensors, what foreseeable risks, patients' privacy concern, or discomforts may occur in using the sensors, and intended research purposes and data use.

There are other advantages of using multi-modal data to improve clinical studies. For example, with the technological advancement and widespread adoption of consumer grade wearable devices that contain multiple built-in sensors, we can obtain individualized baseline

that is close to "truth" using patients' historical wearable device data to assess change in digital measures over weeks, months, or even years. Similarly, this historical device data can help prescreen patients for particular phenotypes and characteristics of interest to select patient cohorts for early phase (I or II) studies. This has the potential to improve the efficiency of these typically small studies by reducing variability.

4. Conclusion

Multi-modal sensors are beneficial to clinical studies by providing a holistic picture of human behavior and physiology in real-life. A broader application of advanced methodologies and innovative approaches to analyze data from multi-modal sensors are needed for researchers to fully utilize those valuable data.

5. Reference

Aktaruzzaman, M. *et al*. (2017) 'Performance comparison between wrist and chest actigraphy in combination with heart rate variability for sleep classification', *Computers in Biology and Medicine*, 89, pp. 212–221. doi: 10.1016/j.compbiomed.2017.08.006.

Battelino, T. *et al*. (2019) 'Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range', *Diabetes Care,* 42(8), pp. 1593–1603. doi: 10.2337/dci19-0028.

Chen, L. *et al*. (2021) 'Deep learning based multimodal complex human activity recognition using wearable devices', *Applied Intelligence,* 51(6), pp. 4029–4042. doi: 10.1007/s10489-020-02005-7.

Cornelissen, G. (2014) 'Cosinor-based rhythmometry', *Theoretical Biology and Medical Modelling*. BioMed Central, 11(1), p. 16.

Di, J. *et al*. (2019) 'Joint and Individual Representation of Domains of Physical Activity, Sleep, and Circadian Rhythmicity', *Statistics in Biosciences,* 11(2), pp. 371–402. doi: 10.1007/s12561-019-09236-4.

Di, J. *et al*. (2022) 'Considerations to address missing data when deriving clinical trial endpoints from digital health technologies', *Contemporary Clinical Trials,* 113, p. 106661. doi: 10.1016/j.cct.2021.106661.

Digital Medicine Society (DiME) (2021) *DiMe's Library of Digital Endpoints*. Available at: *https://www.dimesociety.org/communication-education/library-of-digital-endpoints/*.

Eren-Oruklu, M. *et al*. (2009) 'Estimation of Future Glucose Concentrations with Subject-Specific Recursive Linear Models', *Diabetes Technology & Therapeutics,* 11(4), pp. 243–253. doi: 10.1089/dia.2008.0065.

Facchinetti, A., Sparacino, G. and Cobelli, C. (2011) 'Online Denoising Method to Handle Intraindividual Variability of Signal-to-Noise Ratio in Continuous Glucose Monitoring', *IEEE Transactions on Biomedical Engineering,* 58(9), pp. 2664–2671. doi: 10.1109/TBME.2011.2161083.

Gaynanova, I. and Li, G. (2017) 'Structural Learning and Integrative Decomposition of Multi-View Data'. Available at: *http://arxiv.org/abs/1707.06573*.

Georgiev, A. A., Ramsay, J. and Silverman, B. (1998) 'Functional Data Analysis', *Technometrics,* 40(3), p. 260. doi: 10.2307/1271190.

Goldsmith, J. *et al*. (2016) 'New Insights into Activity Patterns in Children, Found Using Functional Data Analyses', *Medicine & Science in Sports & Exercise,* 48(9), pp. 1723–1729. doi: 10.1249/MSS.0000000000000968.

Karas, M. *et al*. (2019) 'Accelerometry Data in Health Research: Challenges and Opportunities', *Statistics in Biosciences,* 11(2), pp. 210–237. doi: 10.1007/s12561-018-9227-2.

Li, G. and Gaynanova, I. (2017) 'A General Framework for Association Analysis of Heterogeneous Data', *arXiv preprint arXiv:1707.06485*.

Lock, E. F. *et al*. (2013) 'Joint and individual variation explained (JIVE) for integrated analysis of multiple data types', *The Annals of Applied Statistics*, 7(1), pp. 523–542. doi: 10.1214/12-AOAS597.

Mahalingaiah, S. *et al*. (2021) 'Design and methods of the Apple Women's Health Study: a digital longitudinal cohort study', *American Journal of Obstetrics and Gynecology. d*oi: 10.1016/j.ajog.2021.09.041.

Marler, M. R. *et al*. (2006) 'The sigmoidally transformed cosine curve: a mathematical model for circadian rhythms with symmetric non-sinusoidal shapes', *Statistics in Medicine,* 25(22), pp. 3893–3904. doi: 10.1002/sim.2466.

Merikangas, K. R. *et al.* (2019) 'Real-time Mobile Monitoring of the Dynamic Associations Among Motor Activity, Energy, Mood, and Sleep in Adults With Bipolar Disorder', *JAMA Psychiatry,* 76(2), p. 190. doi: 10.1001/jamapsychiatry.2018.3546.

Nweke, H. F. *et al.* (2018) 'Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges', *Expert Systems with Applications,* 105, pp. 233–261. doi: 10.1016/j.eswa.2018.03.056.

Prendin, F. *et al*. (2021) 'Forecasting of Glucose Levels and Hypoglycemic Events: Head-to-Head Comparison of Linear and Nonlinear Data-Driven Algorithms Based on Continuous Glucose Monitoring Data Only', *Sensors,* 21(5), p. 1647. doi: 10.3390/s21051647.

Quer, G. *et al.* (2021) 'Wearable sensor data and self-reported symptoms for COVID-19 detection', *Nature Medicine,* 27(1), pp. 73–77. doi: 10.1038/s41591-020-1123-x.

Rabby, M. F. et al. (2021) 'Stacked LSTM based deep recurrent neural network with kalman smoothing for blood glucose prediction', *BMC Medical Informatics and Decision Making,* 21(1), p. 101. doi: 10.1186/s12911-021-01462-5.

Spira, A. P. et al. (2021) 'Brain Amyloid Burden, Sleep, and 24-Hour Rest/Activity Rhythms: Screening Findings from A4 and LEARN', *SLEEP Advances*. doi: 10.1093/sleepadvances/zpab015.

Uddin, M. Z. and Soylu, A. (2021) 'Human activity recognition using wearable sensors, discriminant analysis, and long short-term memory-based neural structured learning', *Scientific Reports,* 11(1), p. 16455. doi: 10.1038/s41598-021-95947-y.

US FDA (2021) *Digital Health Technologies for Remote Data Acquisition in Clinical Investigations.*

Varma, V. R. et al. (2017) 'Total volume of physical activity: TAC, TLAC or TAC($λ$).', *Preventive medicine.* doi: 10.1016/j.ypmed.2017.10.028.

Yuda, E. et al. (2017) 'Sleep Stage Classification by a Combination of Actigraphic and Heart Rate Signals', *Journal of Low Power Electronics and Applications,* 7(4), p. 28. doi: 10.3390/jlpea7040028. ■