Multiscale Digital Twin Models of Health Behaviors and Behavioral Change

Misha Pavel and Holly Jimison, Northeastern University

Human behaviors are emerging to be key determinants of human health and quality of life. Although more than half of premature mortality can be attributed to undesirable health-related behaviors, changing health behaviors has proven to be a considerable societal challenge. Since successful interventions based on experienced human coaches are not economically feasible at scale, our goal is to develop approaches using mobile and wearable technology that could be used for scalable and financially viable interventions. Although recent advances in sensors and information communication opened unprecedented opportunities to investigate new approaches to interventions, the success of these technology-based methods has so far been, with few exceptions, uninspiring.

We hypothesize that these failures are in part the results of inadequate understanding and models of the dynamics of the processes underlying behavior change. Useful digital twin models of individuals must capture the dynamical and multiscale nature of their behaviors. We cast the optimal intervention problem in the framework of control systems theory that requires multiscale dynamical models to infer and predict human behaviors and behavior change. These models are generally represented by latent multiscale state-space models that govern observable behaviors such as physical activities, sedentary behaviors, nutrition, etc. We can infer a subset of the state variables from subjective responses to short ecological momentary assessments (EMA). This framework opens the question of how to select among various models of behaviors.

Currently, the selection of the specific models is somewhat arbitrary and frequently driven by the ability of randomly selected models to fit the data using statistical machine learning procedures. An alternate approach that is based on causal and mechanistic reasoning would guide the model selection based on the types of constraints imposed by the specific models. The advantages of this causal approach are generalization beyond the training set and explanatory transparency of the models. One specific example of an advantage of this approach is the derivation of the required sampling rates and development of approaches for sparse sampling. We will elucidate this approach by characterizing the models along several dimensions including robustness, observability, sampling frequency, controllability, generalization, and set of invariant properties. We will also discuss the transition from data-driven, statistical computing to mechanistic (causal) models. We will provide examples of the implications of various decisions models using real data as well as computer simulations.