Modeling User Behavior in Context: A Systems-Level Approach to Mobile Health

Anna N. Baglione

A dissertation submitted to the Faculty of the

University of Virginia

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the

Department of Engineering Systems and Environment

Afsaneh Doryab, Chair

Laura E. Barnes, Advisor

Philip I. Chow

Tariq Iqbal

Patrick C. Shih

December 2022

Charlottesville, Virginia

Modeling User Behavior in Context: A Systems-Level Approach to Mobile Health

Anna N. Baglione

ABSTRACT

Mobile devices such as smartphones and smartwatches have fundamentally shifted the healthcare landscape towards more individualized care. Advancements in passive sensing over the past decade have enabled consumer mobile devices to conduct unobtrusive, in-the-moment monitoring of objective features of health and wellbeing. Viewed together, these passively-sensed data comprise a system in which a user's context shapes their future health and behavior. In this dissertation, I take a systems-level approach to understanding health in context. To this end, I present our work on mobile sensing for symptom and medication adherence monitoring. I then present COMP-SCT, a novel framework for deriving personal, behavioral, and environmental features of user context at multiple time scales using Social Cognitive Theory. I apply COMP-SCT to two case studies, demonstrating its utility in using personal and behavioral factors such as mood, levels of engagement, and medication-taking behavior to predict mood and medication adherence among breast cancer patients. This work advances the state-of-the-art in translating raw, passively-sensed data into clinically-relevant insights for personal health monitoring in the wild.

Acknowledgments

I am fortunate to have had the support of many wonderful people in my graduate, master's, and undergraduate programs. Firstly I would like to thank my family and friends. Thank you to my parents, Lois and John, who instilled in me a passion for education from a young age. My dad never denied us a trip to the library or bookstore, and my mom drove me many miles over multiple summers to college prep programs. These labors of love did not go unnoticed. Thank you also to Anthony, Dominic, Sonia, Navreet, Josie, Amber, Erin and Family, and SongHui, who supported me and helped me to stay grounded throughout this journey.

Next, thank you to my mentors, collaborators, and colleagues, who helped shape me into the researcher and person I am today. Thank you especially to my PhD advisor, Dr. Laura Barnes, for her unending kindness and mentorship; to my master's program advisor, Dr. Patrick Shih, who never stopped supporting me; and to my committee members, Dr. Afsaneh Doryab, Dr. Philip Chow, and Dr. Tariq Iqbal, who encouraged me to produce my best work. Thank you also to Dr. Bethany Teachman, who acted as a second advisor and helped me to form lasting collaborations and friendships outside of Engineering; to my colleagues in the IU Proactive Health Informatics Lab and in the UVA Sensing Systems for Health Lab; to my collaborators especially Dan Funk, Dr. Mehdi Boukhechba, Dr. Lihua Cai, Dr. Ryan Gentzler, Dr. David Riley, Dr. Richard Hall, Dr. Kristen Wells, Dr. Jiaqi Gong, and Dr. Luke Dahl; and to all my other mentors especially Gary Jorgenson and

the Robintek team, Bill Richards, Dr. Robert Finn, Dr. Stephen David Beck, Dr. Christopher Branton, Dr. Kay Connelly, Mary Ann Rees, Ken Collura, Dr. Stephen Gilbert, Doreen Close, Nikki Strader, Kathleen Harper, Kay Bea Jones, Dr. Paul Sivilotti, Dr. Johanna Devaney, and Robert Goska and the Clinical Informatics team at Evolent Health.

Additionally, thank you to the following organizations and individuals for their generous financial support of my graduate studies: The Jefferson Scholars Foundation, Mr. Peter and Mrs. Crisler Quick, the National Science Foundation, the Emily Couric Cancer Center at the University of Virginia, the National Cancer Institute, and the National Institutes of Health.

Thanks also to Molly, for being there through all the ups and downs of grad school.

Finally, thank you to my husband, Tyler, for his unending love and support. He is truly my better half and my best friend. I could not have done this without him.

Anti-Acknowledgments

I do *not* gratefully the Ohio State University Honors Collegium (now superceded by the Eminence Fellows Program), for discouraging my participation in NSF-sponsored undergraduate research, and for terminating my membership and the memberships of multiple other female STEM students during a particularly challenging undergraduate semester for "lack of engagement". Your biased decision and callous disregard for my commitment to academic rigor over superficial displays of prestige were significant disruptions to my path to graduate school and will soon be the topic of a lengthy letter to the OSU Alumni Association.

Contents

List of Figures

List of Tables

List of Abbreviations

AE adverse event

ALK A type of NSCLC gene rearrangement

AUC area under the curve

BIT Behavioral Intervention Technology model

EGFR A type of NSCLC mutation

eHealth electronic health

EHR electronic health record

EMA Ecological Momentary Assessment

FACT-B Functional Assessment of Cancer Therapy-Breast

GPS Global Positioning System

HCI human-computer interaction

HIPAA Health Insurance Portability and Accountability Act

HITECH Health Information Technology for Economic and Clinical Health Act

HL7 Health Level 7

HRV Heart Rate Variability

JITAI Just-in-Time Adaptive Intervention

LMM Linear Mixed Model

LOSOCV leave-one-subject-out cross-validation

LR Logistic regression

MEMS Medication Event Monitoring System

mHealth Mobile health

ML machine learning

MMI Multiscale Monitoring and Intervention system

MODAL Mhealth OrienteD mAchine Learning

NSCLC non-small cell lung cancer

PRO Patient-Reported Outcome

RF Random Forest

SCT Social Cognitive Theory

SHAP Shapley Additive Explanations

SMOTE Synthetic Minority Oversampling TEchnique

SMS short message service

SVM support-vector machine

TKI tyrosine kinase inhibitor

WHO World Health Organization

XGB XGBoost

Chapter 1

Introduction and related work

1.1 Background and Motivation

Mobile devices are tightly integrated into the fabric of our fast-paced, on-the-go lives. Over 85% of people worldwide own a smartphone as of 2021 [1], according to the Pew Research Center. "Smart devices", which are easily accessible and affordable and which come with a plethora of native sensors, can act as digital mirrors into our well-being by continuously and passively tracking physiological, social, and physical indicators of our health.

Mobile health, or "mHealth", is use of mobile technologies for health monitoring in everyday life. This rapidly-growing field has revolutionized the healthcare landscape over the past few decades by delivering low-cost, personalized interventions to individuals regardless of their location or ability to receive regular in-person care. Methodological advancements in processing and analyzing mHealth datasets have led to a deeper understanding of how symptoms of illness manifest in daily life [2, 3, 4, 5], and have opened the door to personalized interventions delivered at the very moment they are most needed [6].Heart rate variability can be used to approximate an

individual's level of stress or physical exertion [4], and audio from voice recordings can be used to detect positive or negative emotion [7]. Passively-sensed data have also been used to track disease outcomes and progression by detecting subtle changes in an individual's behavior over time. Changes in movement patterns as detected by GPS, for instance, have been linked to depression [8, 9] and social anxiety [10], and language patterns in both private and public communications have been linked to mental illnesses, broadly [2, 3, 11].

Despite these advancements, qualitative and quantitative approaches to mHealth intervention development remain largely siloed, such that advances in the modeling and prediction of clinically-relevant indicators of health remain divorced from the design and deployment of mHealth tools for personal health management in daily life. Moreover, many mHealth applications lack an explicit theoretical grounding in behavioral theory [12, 13, 14, 15]. These limitations have led to staggering rates of dropout [16] for mHealth interventions in the wild and major gaps in the field's understanding of the the relationship between user context and broader patterns of behavior.

1.2 Dissertation Overview

In this dissertation, I present a systems-level approach to mHealth that focuses on modeling user behavior in context. I ground my approach in Social Cognitive Theory (SCT), which reasons that *personal*, *behavioral*, and *environmental* factors (i.e., constructs) all work together to form the context of an individual's health and wellbeing.

Through a series of studies, I demonstrate the utility of SCT construct monitoring and prediction in oncology contexts.

In this chapter, I review the relevant literature on user context, including measurements of context via digital phenotyping and the relationship between context and engagement. I also present relevant related works in predicting medication adherence from passively-sensed data using machine learning, which is the primary focus of my thesis work.

In **Chapter 2**, I discuss our work on the Multiscale Monitoring and Intervention (MMI) system for guiding the design and deployment of mHealth interventions for medication adherence.

In **Chapter 3**, I present our qualitative study of lung cancer patients' perceptions of and needs for adherence monitoring interventions.

In **Chapter 4**, I present COMP-SCT, a computational framework for Social Cognitive Theory. I then apply COMP-SCT to two case studies of breast cancer patients and demonstrate its utility for predicting personal and behavioral SCT constructs.

Finally, **in Chapter 5**, I summarize my contributions and provide a path forward for future works in mHealth for user context detection.

1.3 Understanding User Context

1.3.1 Defining "Context"

Humans are complex beings. Our thoughts and feelings, immediate environment, and skills and tendencies all collectively work together to influence our everyday behaviors. In order to understand the behaviors a user exhibits during an mHealth intervention (and, subsequently, the effectiveness of the intervention), researchers must first have a firm understanding of the user's *context*. *Context* is both multi-faceted and complex, and researchers have historically taken differing views on what exactly "context" comprises. In the thick of the Dot Com era, when computing devices and personal sensors were rapidly gaining popularity, Salber et al. desribed context as environmental factors that could be sensed: "location, identity, activity, and state of people, groups, and objects". This definition, though perhaps narrow in scope, reflected a paradigm shift in the field of ubiquitous computing (a parent field of mobile health) towards personalized sensing. Dourish later envisioned context as a more abstract phenomenon, calling it "an emergent property of occasions of interaction, rather than being a stable, objective set of features that externally characterize activity" [17].

Theories of health behavior have also alluded to the meaning of context in differing ways. The Health Belief Model [18] highlights how a person's feelings and perceptions, as well as internal and external stimuli (e.g., symptoms, advice from trusted friends or family), form the context that influences a person's beliefs about their

disease risk and recommended care. Similarly, the Theory of Planned Behavior positions context as the collection of an individual's beliefs about factors such as their own behavior and subjective norms, which work together to influence a person's perceived control and intention to engage in certain behaviors [19]. In this dissertation, I draw upon the *Social Cognitive Theory (SCT)* to define and describe user context. SCT positions context a system of personal, behavioral, and environmental influences that, together, shape a person's health and wellbeing.

Importantly, context is both *personal* and *temporal*. It is personal because it is highly dependent on the individual; no two people can have the same context. Moreover, it is temporal because it changes over time; one's context in the present cannot be the same as it was before. Consider, for instance, a young woman who wears a smartwatch as she meets a friend at a coffee shop to catch up. Context in this scenario includes not only the table at which they sit or the kind of coffee they order, but also the women's prior experiences and internalized representations of the world. The smell of a particular coffee blend may evoke pleasant memories, while the clang of a tray crashing to the floor may evoke a fear response. In either of these scenarios, the smartwatch might detect a change in heart rate based on the *inner* and the *outer* facets of context at play. Moreover, the woman's context will continue to change throughout the course of the coffee chat. If she moves from her chair to a different table, her location changes; if her friend brings up an unpleasant topic, her skin may become flushed and hot. These are just a few examples of how context is both personal and temporal.

1.3.2 Digital Phenotyping for User Context Detection

Salber et al. described "context aware" applications as those which use context, often an orchestration of multiple information sources, to modify the behavior of the application [20]. In the case of modern mobile sensing, the multiple information sources used to construct context include the native sensors found in any smartphone, including GPS, accelerometer, gyroscope, microphone, and the circuitry that captures call and text logs. The growth of ubiquitous computing, coupled the rise in popularity and affordability of personal sensing devices, have led to an important subfield of mHealth focused entirely on harnessing user context: *digital phenotyping*. Torous et al. define *digital phenotyping* as "the moment-by-moment quantification of the individual-level human phenotype in-situ using data from personal digital devices" [21]. Over the past decade, digital phenotyping has become the gold standard for context detection and representation in mHealth interventions. A person's digital phenotype is a depiction of their daily physiological functioning as inferred from passive sensor data streams such as GPS, and accelerometer data, as well as from user-supplied data streams such as *ecological momentary assessment* (EMA; [22]). At its core, digital phenotyping aims to quantify what Nahum-Shani, Hekler, and Spruijt-Metz identified as the *static* and *dynamic (or time-varying)* components of an individual's context. For instance, an person's heart rate may rapidly rise as he begins a brisk jog, then gradually slow again as he winds down from his workout. This change in heart rate is part of the individual's dynamic context. In contrast, factors such as his sex (male) remain unchanged during physical activity; this is his *static* context.

Successful digital phenotyping begins with gathering *clinically-relevant* data, which includes all data that can be used to understand more about a person's physiological, mental, emotional, and social wellbeing. Measures of physiology include data gathered through low-level sensor streams via smartphones and smartwatches, such as heart rate, skin temperature, and location and movement (as measured by GPS, accelerometer, and gyroscope). Measures of mental and emotional wellbeing typically include mood scores (e.g. anxiety or depression), personality traits (e.g. neuroticism or openness), and even coping strategies (e.g. dealing with medication side effects); these tend to be administered at baseline, post-study, and at regular intervals throughout the course of a study. Finally, measures of socialization and social wellbeing include audio data captured by a microphone, location and movement data, and call and text logs. Consider, for example, an individual exhibiting signs of loneliness. The microphone on her phone can pick up the sound of her talking with someone else, and call and text logs provide hints that she is communicating with others at certain times [5].

1.3.3 A Systems-Level View of Context Detection

Early successes in capturing user context via digital phenotyping represented a collective watershed moment for the relatively young field of mobile health. On the one hand, the ease with which technology could be leveraged to produce a more granular representation of human behavior was a promising step toward behavior change interventions. On the other hand, existing theories of behavior change could not ade-

quately explain how humans might react to the immediate "nudges" or feedback that mobile devices can provide [23]. This tug of war between older theories and newer technologies has led to the development of various frameworks for user context detection. mHealth frameworks are designed to guide the development of interventions and vary in both scope and purpose. One of the early well-known frameworks for user context representation is Salber et al.'s Context Toolkit, which introduced the concept of "context widgets" that abstract the low-level complexity of sensors in a modular way [20]. Importantly, the Context Toolkit laid the foundation for modern sensing protocols for activity, location, and identity sensing. Other frameworks soon followed, with some such as Mohr et al.'s Behavioral Intervention Technology (BIT) model [24] taking a more theoretical approach. The BIT ties theoretical aims and behavior change theories to practical implementations of behavior-change interventions. The BIT model specifically laid the foundation for how behavior change interventions that leverage digital phenotyping might collect and transfer data between data storage containers (e.g., databases), intervention planners (e.g., algorithms for mapping data to interventions), and the user interface (e.g., a mobile phone screen). Other notable comprehensive frameworks include Cheng et al.'s OmniSense framework for context detection using mobile phones [25], Chow et al.'s DEMONS framework for physiology and self-reported affect [26], and Doryab et al.'s generalizable framework for feature extraction from time series data [27]. Notable contributions have also been made in area of systems development, which focuses on the mechanisms that facilitate the capture and processing of raw sensor data. These include the Ohmage participatory sensing system [28, 29], Ferreira et al.'s AWARE framework for mobile

context instrumentation [30], Chow et al.'s opportunistic and participatory sensing system for social anxiety [31], the Sensus mobile sensing system [32], and the mCerebrum platform for digital phenotyping [33].

1.3.4 Open Challenges

Battery Life

One of the prominent challenges across studies of human behavior using digital mobile devices is battery life. The relatively short battery life of many consumer-grade wearables [30, 34] and the battery drain caused by mobile sensing platforms [35] can cause frustration for the user, as they are required to recharge their device more frequently which, in turn, disrupts their daily routine [36]. Moreover, these power consumption issues can cause problems for researchers, leading to missing data when a device suddenly shuts off for lack of battery power, or when the user has forgotten to charge the device [6, 36]. Recent works have proposed solutions for smartwatch battery drain, including changing the color of the display to be more energy efficient and optimizing CPU usage based on workload [34]; such strategies could likely be extended to smartphones. Ideally, sensor collection platforms should be built from the ground up with battery and performance considerations in mind, so that platforms like Ferreira et al.'s AWARE framework [30] (which did not significantly increase battery drain) can become the norm.

User Burden

User-initiated content is a foundational building block for context representation within mHealth systems [30]. Many context-detection studies rely on frequent user input collected via *ecological momentary assessment (EMA)* [22], an approach to capturing in-the-moment measures of user context and behavior [37]. EMA is generally preferred as a data-gathering strategy because it avoids the *recall bias* from which more traditional retrospective analyses suffer and because it cultivates feature granularity. EMA questions are often delivered as surveys to a smartphone or wearable device, either at fixed or random intervals throughout the day. Unlike the native sensors in smartphones and smart watches, which passively capture data without requiring user input, EMA requires user response to capture a meaningful depiction of user context, and thus frequently interrupt the user's daily routine. The influx of notifications from EMAs renders mHealth devices vehicles a major source of user fatigue [38] and may place an undue burden on the user over long periods of time. This prolonged burden can lead to attrition [39, 40] and, consequentially, less data from which to learn user behavioral patterns. Consider a machine learning classifier that trains only electronic health record (EHR) data. This classifier might suffer low accuracy rates without information provided by the user, such as family history of disease or date of diagnosis of depression, as these factors are of critical import when searching for factors that influence both disease and treatment outcomes. Similarly, mHealth applications will lack a fundamental understanding of context without some form of user-inputted data, such as one's baseline depression score or "home" GPS location; the user must *become* a context sensor, in some sense,

to generate meaningful data [30].

1.4 Understanding User Engagement with mHealth interventions

This section is adapted from forthcoming work conducted with Dr. Sonia Baee and colleagues on attrition prediction in electronic health (eHealth) studies.

1.4.1 Quantifying Engagement and Attrition

A key part of understanding an individual's context is understanding their actions. User engagement, in the mHealth literature, describes an individual's interactions with a given intervention and their level of motivation during these interactions [41]. Measures of engagement captured by passive sensing devices can reveal much about a user's context, such as the time of day at which a user is most receptive to changing their behavior. Measures of engagement in mHealth interventions include, but are not limited to, the number of questions completed, number of total clicks on the intervention content, time spent on each page, and time at which the user completes a questionnaire. Recent works have found success in extracting behavioral features from log of usage data, including duration [42, 43, 44], number of actions [42, 44], duration per session [44] and duration between sessions [45]. Sustained engagement with mobile interventions over the course of treatment is a common and persistent problem [16, 46]. In his seminal 2005 paper, Eysenbach characterized the phenomenon of

attrition, which occurs when participants either leave the study prematurely or are lost to followup [16]. Attrition is inherently linked to engagement, such that interventions which fail to properly sustain the user's interest and attention may lead to the participant giving up on the intervention entirely. Attrition is a particular threat to mHealth interventions in that it can bias the results and limit the generalizability of intervention studies, given that only a small, potentially homogeneous participant population remains. Researchers have therefore invested a great amount of time into quantifying attrition, with the goal of preventing it be fore it happens. Approaches to attrition analysis remain divided between academic disciplines. Among psychologists and psychiatrists, the early pioneers of eHealth interventions, attrition analysis is primarily comprised of statistical modeling using techniques such as ANOVA and regression [47, 48, 49]. Researchers within computer and data science and outside academia, in the mobile gaming industry, have found recent success in predicting attrition ("churn") using more advanced techniques such as linear mixed modeling [50], survival analysis [42], probabilistic latent variable modeling [51], and neural networks [42, 43, 52, 53, 54].

1.4.2 Factors Which Influence Engagement and Attrition

Both demographic and behavioral factors have been shown to influence engagement and attrition in mHealth interventions. In particular, age, gender, race, and education have been shown to be predictive of attrition [49, 55, 56]. Ben-Zeev et al.'s evaluation of the FOCUS program for schizophrenia found that younger, white, fe-

male participants were the most engaged with the intervention [57]. Further, Blixen et al.'s focus group study with individuals with comorbid bipolar disorder and hypertension uncovered common issues associated with both medication nonadherence and lack of engagement with the SMS-based intervention [58]. These factors included forgetfulness (most common), being very busy, concerns about side effects, impaired decision-making due to BD, and the feeling of having too many medications to manage. Notably, these studies highlighted the potential for decreased engagement among populations that have suffered systemic disadvantages, such as lack of access to quality medical care and safe, affordable housing.

Intervention design also plays an important role in engagement and attrition. Design choices such as the level of personalization in the intervention, the user's perception of the intervention's effectiveness, and the availability of human support during the intervention [59] all affect whether a user is likely to drop out. Studies aimed at attrition reduction have also explored the effectiveness of *rotation* of intervention components, in which different pieces of the intervention are cycled out to provide variety to the user. In a 2018 study of a web browser designed to help users take better control over the time they spent online, the authors found that rotation functioned as a "double-edged sword", increasing both engagement and attrition [50]. However, including components such as a popup warning of the next rotation helped mitigate the overall negative impact on attrition.

Patients have also identified several key factors that impact their engagement, the first of which is *patient-provider communication*. Brath et al. examined the use of the medication adherence measurement system (mAMS) and a telehealth system

with short message service (SMS) reminders among 150 older adults with comorbid cardiovascular risk [60]. Participants appreciated when their doctor was attuned to their adherence or nonadherence, indicating that closer direct (or perceived) involvement from doctors may help increase engagement. Similar calls for increased provider involvement, such as via two-way patient-provider communication channels, were raised by Chiang et al's cohort of patients using the WelTel antiretroviral therapy platform [61] and by Hilliard et. al.'s cohort of cystic fibrosis patients [62]. Another factor participants value is *social support* from family and friends, who help remind patients to take their medication regularly and help them to stay on track with their treatment [58]. Indeed, patients have requested that additional resources for accessing social support be included in mHealth applications [38, 62]. Finally, participants value *control*, such as the ability to customize and personalize content. Blixen et al.'s cohort indicated that they would be more likely to engage with the intervention (and thus adhere to their medications) if it allowed them to personalize the content sent to them via SMS [58]. Likewise, Guo et al.'s cohort desired feedback personalized to individuals as well as to subgroups of persons living with HIV [38]. These calls for personalization point to a need to explore the impact of partial user control on adherence and engagement, as well as factors like self-efficacy. Kannisto et al.'s work using Mobile.Net with individuals with serious mental health issues is one of a few mHealth studies that allowed for partial user control over the intervention [63]. Participants were allowed to help customize the timing and number of message reminders sent to them, as well as the content of the messages themselves. Interestingly, attrition was quite low for intervention (4.8%), though the

authors note this may be to underreporting of dropout by participants. Regardless, this study presents interesting implications for the interplay between personalization, engagement, efficacy, and attrition.

1.5 Modeling User Behavior in Context: The Case of Medication Adherence

1.5.1 Measuring and Monitoring Adherence

With the increasing threat of chronic diseases such as cancer to the global population, increased attention has been dedicated to studying *medication adherence* and its associated behavioral context in a variety of patient populations. Medication adherence is "the extent to which a person's behavior corresponds with agreed upon recommendations from a health care provider" [64]. Adherence rates among cancer patients tend to vary widely; among breast cancer patients, for instance, a systematic review found adherence ranges from approximately 40% to 70% [65]. This variation can be explained, in part, by the multitude of factors that impact adherence. These include self-efficacy, socioeconomic status, depression or mood, cognitive or physical functioning, side effect severity, cost of care, beliefs in the importance of the medication, communication with clinicians, and social support [66, 67, 68, 69]. The definition of adherence varies from study to study [67] and is therefore difficult to measure. Researchers and clinicians are increasingly seeking more accurate ways to

measure adherence rates among patients taking oral anticancer medications at home, as well as ways to monitor patients for possible side effects or adverse events.

Previous studies have primarily measured adherence via traditional methods such as blood tests or self-reports [70]. More recent studies, however, have incorporated the use of Medication Event Monitoring Systems (MEMS), which can detect when the user has taken their medication [71, 72] MEMS offer several advantages over traditional methods: they are non-invasive, do not encourage recall bias, and reduce intervention fatigue by fitting into the person's daily life, and are a valuable tool for measuring adherence over time. While early MEMS devices tended to be simple sensor-equipped pill caps, the definition of "MEMS" has broadened significantly to include Bluetooth-connected pill bottles and boxes [73], wearable sensors [74], mobile apps [75, 76, 77] ingestible sensors [78], and even entire systems of mHealth devices. Modern MEMS technologies can be used to passively track not only patients' adherence, but also related factors, such as symptoms and problems during treatment [79]. These devices can easily capture*patient reported outcomes* (PROs) such as sleep quality, medication side effects, and mental wellbeing in real time and may lead to better overall outcomes for patients whose physicians closely monitor their reported progress [80]

1.5.2 Predicting Adherence

Machine learning (ML) has revolutionized the process of transforming passively sensed data into actionable predictors of human health and behavior. ML models can

process enormous amounts of complex, multimodal behavioral data in a relatively short amount of time, learn important patterns from this data, and make predictions about a patient's future outcomes based on those patterns. Unsurprisingly, generating predictions using ML models comprises a major portion of digital phenotyping research.

Recent advances in machine learning have enabled large-scale studies of adherence over time with little burden on the patient. Adherence prediction studies commonly use *supervised learning*, a form of machine learning which learns to predict outcomes for labelled sample data based on the qualities (features) of the data. Supervised learning models take, as their input, a dataset D containing labeled, i.i.d. vectors:

$$D = ((x_1, y_1), ...(x_m, y_m)); x \in X, y \in Y \text{ [81]}$$

where X is the input space and Y is the output space. The model learns a mapping $f : X \rightarrow Y$ from D. Supervised learning can be performed as a *regression*, in which the model predicts a continuous outcome (a probability), or *classification*, in which the model predicts a binary outcome. In the case of medication adherence, a regression task might predict the probability that a patient will adhere to their medication, while a classification might simply predict whether they will be adherent or not.

A number of recent studies have leveraged mobile sensing and machine learning techniques such as supervised learning to reveal important findings about the context(s) of medication-taking and the nature of adherence prediction. Accelerometer data,

for instance, has been shown to be predictive of adherence-related actions such as removing a pill from a blister pack [82], placing the pill in one's mouth [82], and swallowing a pill [74].

Studies of human behavior such as medication adherence have shown the importance of using features of prior behavior to predict future behaviors. Yet, time series data are structured to contain one observation per unit of time. Incorporating prior measures necessitates a change to the structure of time series data prior to feeding this data into a supervised learning pipeline. One popular restructuring technique is the *sliding window* method, which constructs one vector per unit of time (e.g., day, week) for which we want to predict our outcome. Each vector contains the values of the outcome we want to predict for the previous N units of time, where N is selected *apriori*; these are called *lagged* features. Additional lagged features which are not directly related to the output can also be included. In the case of predicting future medication adherence, for example, we might be interested in prior values of adherence as well as prior self-reported mood.

As a simplified example [1], let us formulate a univariate time series prediction task with dataset D. Let us call our output space D'. For each vector $v_t = (x_t, a_{t+1})$ in D, input x_t contains exactly one variable: a_t, the patient's adherence at time t. Thus we want to predict output a_{t+1}, the patient's adherence at the next time step. Further, let us select a window w, and let us assume is optimal. For each vector v_t, we will generate a new "windowed" vector v'_t by prepending prior adherence values from the

[1] Adapted from examples by Jason Brownlee: https://machinelearningmastery.com/time-series-forecasting-supervised-learning/ and Pablo Ruiz: https://towardsdatascience.com/ml-approaches-for-time-series-4d44722e48fe

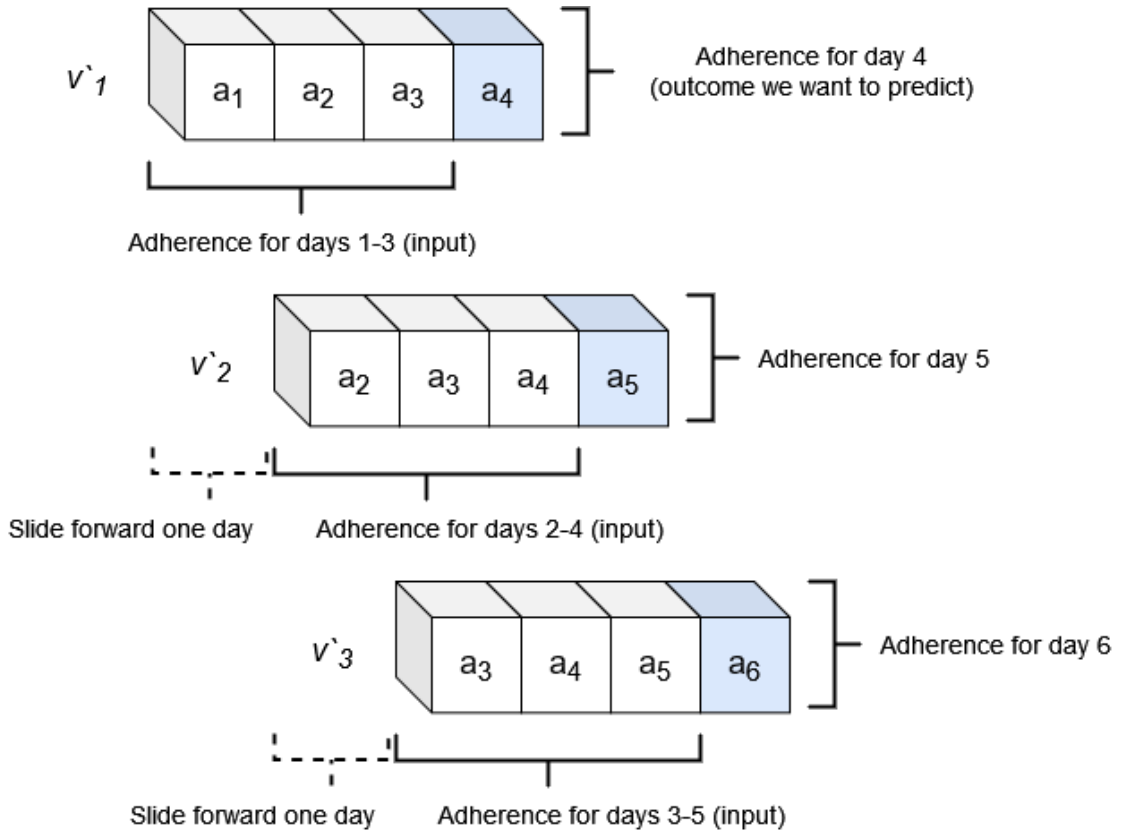


Figure 1.1: Sliding Window Algorithm Visualization

most recent $w-1$ time steps to the original vector; vectors for which $w-1$ prior values of adherence are unavailable are dropped from D. This approach leads to a "sliding window" effect because, for each new vector, we have effectively shifted our view of the previous adherence data forward by one time step. Figure 1.1 provides a visualization of a simplified example of the sliding window method for a window of size $w = 3$ and six total days of adherence data. For further reading on similar approaches, I refer the reader to [83].

Studies which have utilized *lagged* input features have been particularly successful in

uncovering relationships between prior adherence and future adherence at different timescales. Koesmahargyo et al., who sought to predict nonadherence from recorded videos of patients taking their medication, showed that more recent data (e.g., data from the prior day of the study) was more predictive of future adherence than less recent data (e.g., data from the first 7 days of the study) [84]. Importantly, authors also demonstrated an inverse relationship between the number of input features and the false positive rate, and showed that predictive performance is improved with more granular prediction horizons. In other words, leveraging more information about previous adherence can lead to better predictive outcomes, and predicting adherence in the near future (e.g., next day adherence or next week adherence) is more feasible than predicting adherence for the entire remainder of the study. In a similar vein, Gu et al. [85] and Gu et al. [86] conducted two related studies on predicting adherence to injectable medications. The authors not only confirmed Koesmaharygo et al.'s findings that prior adherence is highly important for predicting future adherence, but also made several novel advances in the areas of feature selection and lagged input selection. Namely, they showed that tree-based classifiers such as Random Forest (RF) may be useful for finding the optimal number of lagged inputs and most predictive features, and that combining input optimization, feature preselection, undersampling, and gridsearch with cross-validation may lead to excellent predictive performance (>0.8 AUC) provided that enough samples are used.

Importantly, more advanced models are not always required to achieve good predictive performance in adherence prediction tasks. *Tree-based* models such as Random Forest tend to perform quite well in practice [87], sometimes even outperforming

more complex models such as neural networks [88, 89].

Chapter 2

Leveraging Mobile Sensing to Improve Medication Adherence

2.1 Introduction

Medication adherence is defined by the World Health Organization (**WHO**) as "the extent to which a person's behavior (taking a medicine), corresponds with agreed upon recommendations from a health care provider" [64]. Adherence to long term therapy for chronic illness in developed countries is about 50% [64]. Interventions aimed at increasing medication adherence have the potential to provide a significant benefit through both primary prevention of disease risk factors and secondary prevention of adverse health outcomes. In fact, increasing the effectiveness of medication adherence interventions may have a far greater impact on health outcomes than any improvement in specific medical treatments [90]. The impact of poor adherence to medications is expected to continue to increase as the burden of chronic

disease increases globally. Endocrine therapy, including aromatase inhibitors and Tamoxifen, is prescribed for at least 5 years after individuals have been treated for hormone receptor-positive breast cancer to prevent recurrence of their cancer. Adherence to these medications (defined as 80% or more doses taken as prescribed) is associated with significant increases in recurrence-free survival [91]. Despite the life-saving benefits of these medications, rates of persistence and adherence are low [91], with post-treatment adherence ranging from 41% to 72% and discontinuation ranging from 31% to 73% [65].

Increasing adoption of smartphones has led to an upsurge in applications targeted at improving medication adherence. However, these technologies have focused mainly on cognitive factors contributing to nonadherence, such as managing multiple medications [92]. Existing mobile applications fail to account for an individual's specific risks and fail to personalize the interventions delivered according to those risks. Furthermore, very few interventions have been assessed for efficacy in supporting adherence [93].

2.1.1 Contributions

In this chapter, we propose a new integrated system for long-term monitoring of medication adherence consisting of sensor-rich smartphones, wireless medication event monitoring systems (**MEMS**), wireless beacons, and wearable sensors that collect *in situ* data on adherence. This data will be used to understand and model medication-taking behaviors, develop context-sensitive models to predict nonadherence, and de-

velop and deliver personalized interventions to improve medication adherence. The novelty of this project lies in its capturing of the multidimensional complexities of medication adherence using ubiquitous mobile sensing technologies and in using these sensed data to understand medication-taking behaviors, predict individual risk factors, and design and deliver interventions to improve adherence at the optimal time and in the optimal context.

This chapter is based on our work on modeling and intervention for medication adherence in context [94].

2.2 Background and Motivation

2.2.1 Defining and Measuring Medication Adherence

Medication adherence is defined as whether patients take their medication as prescribed [70] and is critically important for effective medical treatment. Methods for assessing medication adherence are categorized as either *direct* (e.g. directly measurement of medicine or biomarkers in blood) or *indirect* [95] (e.g. patient self-report, pill counts, and pharmacy refills [70]). Indirect methods, in particular, have several notable limitations. Self-reports are often biased by inaccurate patient recall or social desirability. Pill counts, meanwhile, do not accurately capture exact timing of medication-taking and can be easily manipulated by patients (e.g., pill dumping). When used alone, these methods fail to provide a deeper contextual understanding of reasons for medication adherence or nonadherence.

Interventions to Increase Medication Adherence

A 2014 systematic review identified 17 randomized controlled trials that evaluated the efficacy of medication adherence interventions [90]. Only five of these studies found medication adherence interventions were associated with both increased adherence and better clinical outcomes, even though most interventions studied were complex and required significant time of healthcare staff. This systematic review concluded interventions may not have been effective because there is a lack of understanding of barriers to adherence and *the context in which adherence and non-adherence occurs*. Another review of 229 smartphone reminder applications (apps) determined that a "one size fits all" timer-based reminder was largely ineffective because it did not consider a user's routine [75]. Taken together, these reviews indicate that previous interventions to increase medication adherence have broadly targeted factors associated with lack of adherence across groups of individuals. However, these intervention approaches may not be relevant to a specific individual at the time and in the context that it is delivered and are not sustainable because of the high burden on health care providers and the healthcare system.

2.2.2 Factors Associated with Medication Adherence and Non-adherence

Substantial research documents reasons why individuals do not adhere to prescribed medications, including endocrine therapy, a life saving medication taken by some

cancer survivors to slow or stop cancer growth. We examine factors associated with medication adherence and nonadherence through the lens of Social Cognitive Theory (**SCT**) [96], a commonly used health behavior theory which can facilitate better understanding of the context of medication taking by evaluating how environmental factors, personal factors, and a person's behavior interact. The SCT guides the selection of constructs in our new medication adherence framework and will also guide intervention development.

Personal Factors: *Physiological*, *cognitive*, and *affective* states affect long-term adherence to medications. Physiological factors significantly associated with endocrine therapy nonadherence include side and adverse effects and functional impairment [65, 91]. The two strongest cognitive predictors of adherence, generally, and endocrine therapy, specifically, are self-efficacy and positive beliefs regarding the importance and necessity of medications [91]. In addition, individuals who are poorly informed about side effects are less likely to adhere to any medication, including endocrine therapy, and that updating patients' knowledge regularly can improve adherence. [97]. Affective states associated with lack of adherence to medications are distress, depression, and fear of cancer recurrence [65, 91].

Environmental Factors: A person's *social* environment, *physical* environment, and the *health system* environment influence whether or not an individual takes medication. Having less than desired social support for taking the medication is linked to endocrine therapy nonadherence [65, 91]. Medication adherence is associated with positive interactions with health care providers who provide medication reminders [91]. Family members also facilitate medication adherence through re-

minders [91]. Medication adherence can also be facilitated by aspects of the physical environment; people frequently place medications in places where they frequently go so that they remember to take them [91]. Health system environments can also affect medication adherence, such as via costs of medication [65, 91].

Behavioral Factors: Medication-taking occurs in the context of other behaviors which can serve as a cue to action to initiate the behavior of interest (e.g., brushing teeth, eating breakfast) [98]. Having a routine or schedule for medication-taking may facilitate adherence [91]. To date, the field has been limited in monitoring multiple factors simultaneously due to limitations in technologies to collect the data and model these factors dynamically. However, the emergence of mobile technologies enabling remote health monitoring and studying human behavioral dynamics [32, 99]. An understanding of the interaction of environmental, personal, and behavioral factors associated with medication-taking in each individual will enable the development of personalized approaches to prevent medication nonadherence [100].

Mobile Sensing and Modeling

Smartphones and wearable technologies have arrays of embedded sensors that measure mobility, location, acoustics, and ambient light. These sensors can be harnessed to passively capture information related to users' personal and environmental factors and behaviors, so long as individuals carry or wear the devices. These technologies are modernizing patient care with capabilities such as sending and receiving clinically-relevant messages and supporting illness management and treatment applications.

Most approaches require individuals to actively engage with the device by responding to prompts or launching an app. However, smart devices and remote sensing technologies can also facilitate behavioral tracking techniques that require little to no active response from the user, thus *decreasing patient burden*.

Mobile behavioral sensing has been used to draw inferences about how and where individuals spend their day and to track behaviors associated with stress and changes in mental health over time. Ben-Zeev et al. used a mobile sensing application to gather GPS, activity, and sleep data in tandem with daily stress ratings gathered via EMA. They identified relationships between sensed data such as sleep and activity and changes in stress and mood [101]. Boukhechba et al. showed that sensed features such as location entropy can be used in tandem with social anxiety baseline measures to predict symptom severity [102]. Further, Gong et al. found that increased accelerometer movement is tied to social anxiety symptoms for activities in certain contexts (e.g. when making a phone call) [99]. The methodologies and metrics from such studies are promising, demonstrating that complex human behavior and psychological states can be inferred from multimodal data. However, no current sensing systems have been implemented to provide continuous monitoring within and outside the home to monitor and support long-term medication adherence.

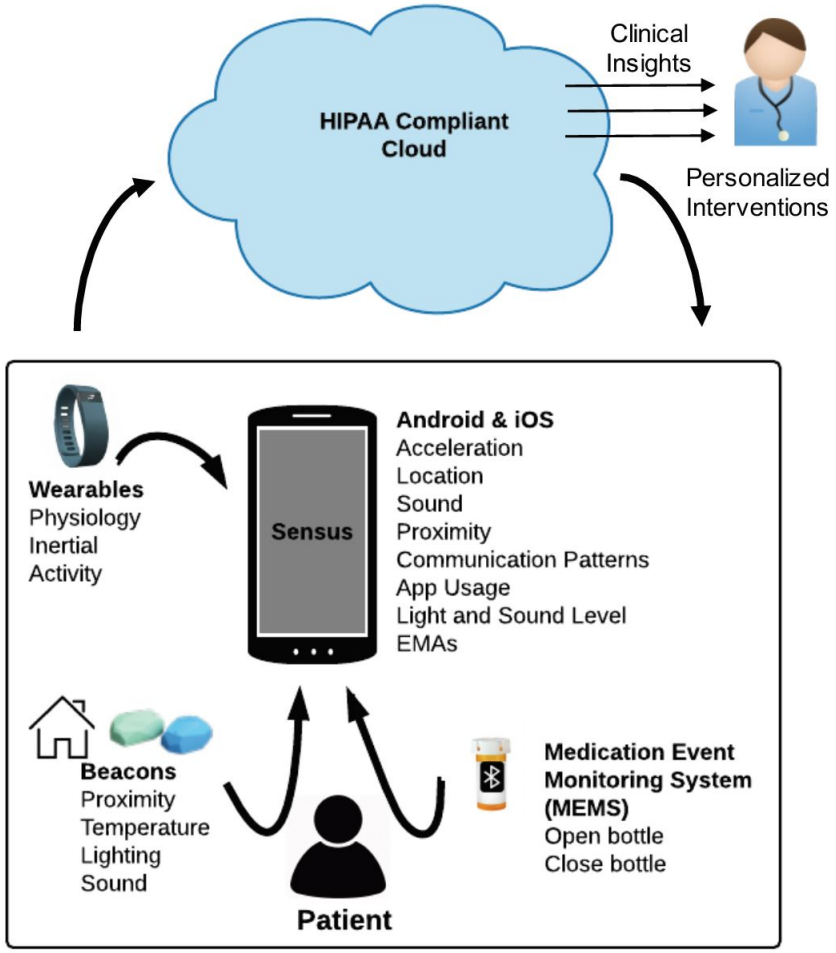


Figure 2.1: MMI sensing framework and data flow overview

2.3 Proposed Framework: Multiscale Modeling and Intervention System (MMI)

Social Cognitive Theory (SCT) [96] indicates medication-taking behaviors are performed in the context of an individual's environments (e.g., work, social) and other behaviors (e.g., eating) and are influenced by personal factors (e.g., cognition, emo-

Table 2.1: SCT Constructs, Measurement Methodology and Exemplar Features

SCT Construct	Data	Sensors	Features
Personal	Hot flashes, rash, pain, fatigue, stress, cognition (e.g. brain fog)	Galvanic skin response (GSR), electrocardiogram (ECG), GPS, accelerometer, gyroscope	Heart rate variability, body temperature, breathing rate, range of movement, times of day for reduced movement
Behavioral	Sleeping, eating, medication-taking, exercising (e.g. walking or running)	ECG, gyroscope, accelerometer, MEMS cap, GPS	Heart rate variability, movement velocity, movement variation, location
Environmental	Social interactions (type / quality / with whom), patient-provider communication	Microphone, GSR, ECG, accelerometer, gyroscope, GPS, app usage logs, call/text logs	Audio signals, social media activity, semantic location diversity, text and call frequency, app use frequency

tion, experiences of side effects). A deeper understanding of the context of medication-taking will provide us with information about how, when, with whom, and where medication is taken. Consequently, we propose a new sensing systems framework, the **Multiscale Modeling and Intervention** (**MMI**) system, for modeling the simultaneous interacting behavioral, environmental, and personal factors that influence medication adherence or nonadherence. We are applying this framework to breast cancer survivors who have completed most treatment and are prescribed long term endocrine therapy. This framework is grounded in SCT and accomplishes four main goals: it 1) senses medication-taking behaviors in context (i.e. sense personal, environmental, and behavioral parameters,) 2) models the complex constructs of medication-taking behavior in context, 3) identifies person-specific constructs and constraints, and 4) establishes a methodological foundation for creating personal-

ized interventions to improve medication-taking behavior. Figure 2.1 provides an overview of the MMI system, and Table 2.1 demonstrates how sensed data map to the three key SCT constructs. We now highlight the MMI system components and design considerations in light of the aforementioned goals.

2.4 System Components

The MMI sensing system unites multimodal sensor data from smartphones, wearable sensors, wireless beacons, and smartphones. In particular, the Sensus adaptable sensing system [32] is central to the MMI system enabling the collection of behavioral data via ecological momentary assessments (EMAs) and smartphone and wearable sensor data. Medication-taking is measured using MEMS devices which do not provide users with feedback regarding previous opening of the bottle. Bluetooth beacons are integrated into the system to transmit environment-specific contextual information to the user's smartphone. Specifically, these beacons are used to sense proximity, temperature, and ambient lighting levels in a participant's significant physical environment locations (e.g. home). These contextual data are then used to learn an individual's event patterns. For example, placing a beacon in an individual's kitchen (e.g., physical environment) will allow the system to learn when they are likely having a meal (e.g. engaging in a common behavior). Passive sensor data from smart devices are coupled with EMAs to collect self-reported ground truth for dynamically changing measures of SCT constructs (e.g., side effects, behaviors related to medication adherence). Participant burden in responding to EMAs is minimized by

keeping assessments brief and leveraging smart-sensing plans to trigger prompts. For example, environmental context (e.g., GPS, beacons) can be leveraged to infer that a participant is eating and prompt the user to confirm.

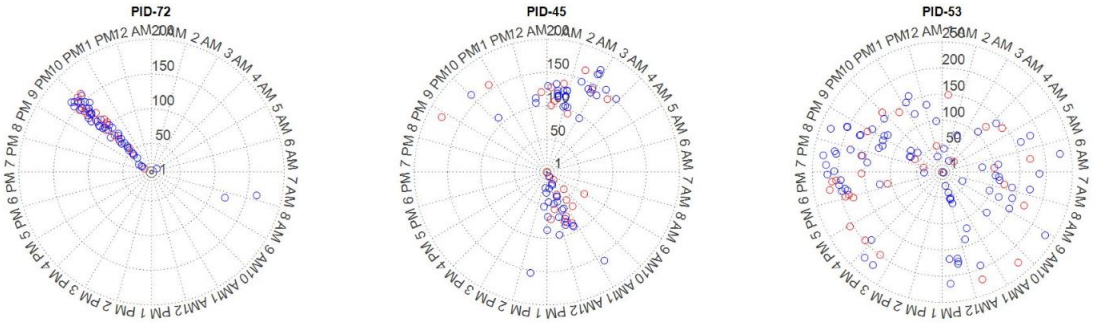


Figure 2.2: A polar coordinates plot demonstrating patterns of medication-taking behavior in breast cancer survivors leveraging MEMS over a eight-month period; consistent evening pattern (patient ID: 72), changing patterns (patient ID: 45) and random pattern (patient ID: 53). Blue dots are the MEMS data in the weekday and red dots present the data in the weekend.

2.5 Design Considerations

2.5.1 User-centered Design

Including users throughout the process of the MMI system design, as well as keeping users engaged with interventions that arise from the design process, is of critical importance. We will conduct participatory design interviews with breast cancer survivors to assess both the usability and acceptability of MMI system components (e.g. wearable sensors, MEMS caps, and EMAs). We will also examine critical markers

of engagement captured by the system, such as dwell time on tasks (e.g removing a MEMS cap or responding to a survey) and task sequences (e.g. responding to a survey, then removing a MEMS cap). These markers are particularly useful for inferring common behavioral constructs and for optimizing intervention timing to maximize engagement. For example, consider a routine such as brushing one's teeth at night. Inferring the time at which this behavior typically occurs (e.g. 8pm) and the behavioral context that is likely to follow (e.g. sleeping) could help the MMI system know when to deliver a reminder to take a medication prescribed to be taken at bedtime. Another example of an engagement strategy the MMI system could employ is a web-based dashboard which displays adherence rates for the past week as well as motivational messages encouraging patients to stick to a medication routine.

2.5.2 Privacy and Security

The MMI system will collect personal and sensitive data, and thus, specific strategies need to be taken to ensure the preservation of user privacy and data security. We will leverage MEMS and wearable devices that offer a secure on-board storage infrastructure and transmission protocols. Participant data will be transferred via APIs that adhere to industry standards (e.g. use TLS encryption and trusted tokens) and will be stored in a HIPAA compliant cloud. We will leverage privacy-preserving data processing methods for potentially identifiable data, such as GPS location, which could be abstracted into clusters to avoid a situation where a user's precise location could be pinpointed.

2.5.3 Energy-Efficient Sensing

One of the critical challenges of mobile sensing research is creating energy efficient systems that minimize power consumption while capturing enough information to measure / predict user context. Adaptive sensing methods can be used to control low-level sensing cycles to collect data only when needed hence minimizing the a device's computational usage (e.g. collect motion data only when the device is moving). Furthermore, machine learning methods can be leveraged to learn when to turn on a specific sensor based on an individual's patterns of daily living.

2.6 Social Cognitive Theory: A Strong Theoretical Foundation

SCT has been used extensively in understanding, predicting, and facilitating adherence to a wide range of behaviors, including medication adherence. The benefit of applying SCT to complex behaviors like medication adherence is that it includes the concept of *reciprocal determinism* and thus considers interactions among environmental, personal, and behavioral factors associated with the medication-taking behavior. We hypothesize not only that the interaction between these factors contributes to medication use, but that a change in one factor may affect other factors to increase or decrease the likelihood of medication adherence. Other health behavior theories, which are less complex, use similar constructs but do not include reciprocal determinism.

Payne et al.'s systematic review showed frequent adaptation of SCT constructs to mHealth interventions, providing evidence of feasibility for our approach [103]. Notably, however, none of the previous works mentioned in this review present a framework that has been developed and validated for the purpose of medication adherence. The MMI system represents the first framework developed by a transdisciplinary team of researchers oriented toward improving medication adherence *in situ*.

Preliminary work from our team has already yielded rich insights into medication-taking behavior patterns and represents a key first step toward building the MMI framework on SCT constructs. Boukhechba et al. analyzed data assessing medication adherence of 33 breast cancer survivors taking endocrine therapy medication using MEMS over an eight-month period [97]. These results indicate that breast cancer survivors have diverse patterns of medication-taking behavior over the course of the monitoring period.

Figure 2.2, provides a visualization of three of these breast cancer survivors' patterns of endocrine therapy medication taking over the study period according to the time of day that each person took her medicine. Participant 72 nearly always took the once daily pill in the evening, participant 45 took the pill in the morning initially and then switched to the evening, and participant 53 took the daily medication at various times of the day with no noticeable pattern. It is important to note that, the MEMS devices used in this study provide no feedback to the user; however, we saw numerous instances of medication non-adherence suggesting that there was not a substantial Hawthorne effect.

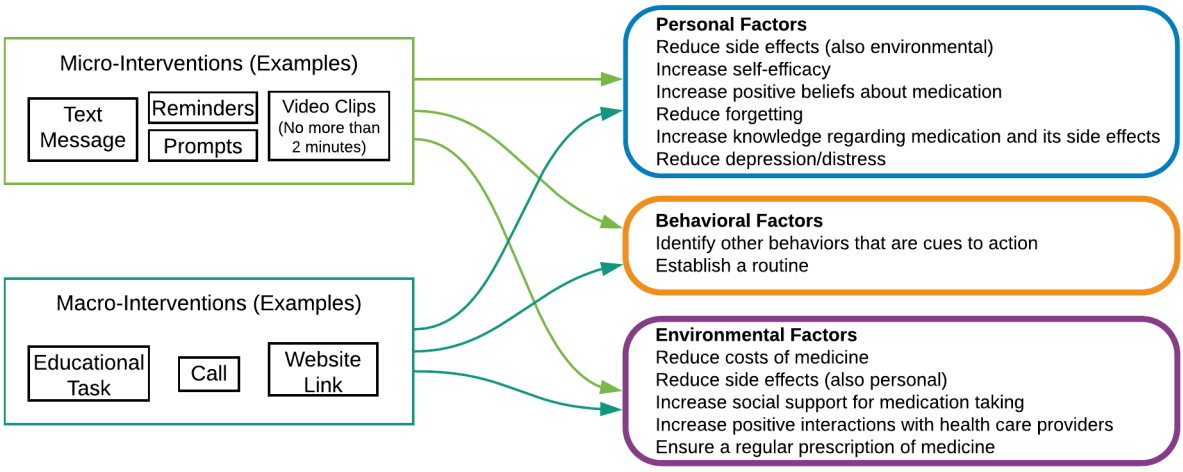


Figure 2.3: Mapping personalized intervention options to contextual factors

2.7 Sensing and Modeling Medication-Taking Behaviors in Context

Medication-taking behavior is part of a person-specific human behavior system and is a system itself. Understanding the complex system of adherence-related context requires comprehension of both stable and dynamic variables. *Static* variables refer to characteristics of one's life which infrequenly or never change, such as personality, intelligence, and some demographics. *Dynamic* variables, on the other hand, refer to frequently-changing characteristics such as medication side effects, disease symptoms, health system interactions, social interactions, and behavioral contexts. In order to capture both static and dynamic variables, the data collected from the MMI system is translated into contextual features within the SCT framework (Figure

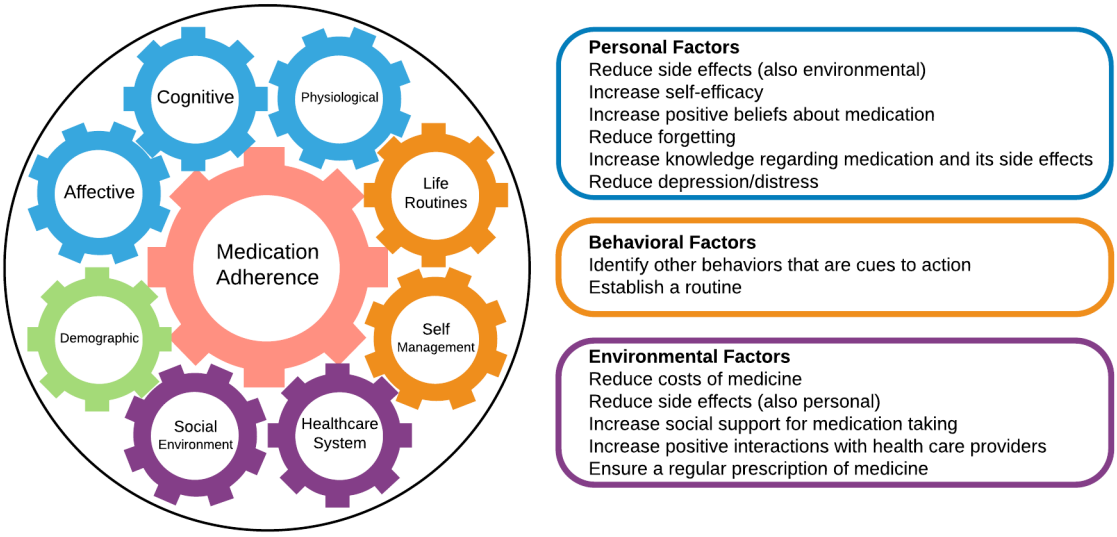


Figure 2.4: Conceptual representation of SCT framework. Interventions will be designed to influence the personal, behavioral, and environmental factors in order to facilitate better medication adherence.

2.4). Machine learning methodologies (e.g., network analysis, hierarchical sampling for active learning) can then be used to discover the complex structure of medication-taking behavior.

2.8 Identifying Person-Specific Constructs and Constraints

Recent works in multiscale pattern recognition has demonstrated that information fusion methods provide richer information than isolated data-driven models [99, 102].

The MMI framework uses pairwise sensor fusion methods at several junctures. For example, fusing information from communication events (identified from call and text logs) and fine-grained motion sensor data yields a reliable behavioral marker of social anxiety levels [99]. High-fidelity time-series data can generate features from sliding windows or change-detected windows, and extracted features can be clustered to identify the semantics of activities. For instance, the relationship between accelerometer data and heart rate data can be examined to understand how stress manifests in daily life. Although the information from an accelerometer sensor is not accurate enough to identify the complex entities of the human activities in daily life (e.g., sleeping, typing), additional integrative models which include heart rate, skin temperature, skin conductance, and other information (e.g., GPS changes, call and text, and EMAs) can be used in combination to more closely approximate activity.

2.9 Development of Personalized Interventions

Development of personalized interventions within the MMI system is achieved via *intervention modules*, which address patient-specific needs or barriers to medication-taking at time and place that is most convenient for the patient. Modules incorporate constructs from SCT and include delivery of brief content (e.g., text, audio, video) on the mobile phone via an app, or on the smartwatch, as well as phone calls and text messages providing personalized content. Interventions are designed to only involve the patient's health care team in addressing barriers that require their assistance, reducing the burden on both patients and health care providers. We note that

frequent low-level interventions, such as reminders, may annoy the users or the users may habituate to them and ignore them. Therefore, determining drawbacks and constraints of contextual factors such as notification fatigue will be a critical step in designing future intervention modules.

Intervention modules are guided by Intervention Mapping [104], an intervention development model used in public health and behavioral sciences. Intervention Mapping includes six steps: logic model of the problem, a logic model of change, program design, program production, program implementation plan, and evaluation. The knowledge learned from the computational models employed in the MMI framework, such as the contextual factors of medication-taking behavior and the constraints of these factors, are uniquely tied to the intervention approaches in the following ways: 1) determining behavioral and environmental outcomes the intervention is targeting (e.g., remembering to take the medication on weekends); 2) stating performance objectives for each outcome; (e.g., realizing that your routine is different on weekends, linking medication to a behavior performed every weekend); and 3) determining essential and changeable determinants using SCT (e.g., determining behaviors that are performed every weekend, placing medication in a visible place after it is taken on Friday) of behavioral and environmental outcomes as indicated in the data collected from the MMI system.

2.10 Deployment of the System with Integrated Personalized Interventions

The natural progression of the development of the MMI system will be deployment with real breast cancer survivors prescribed endocrine therapy. This deployment will enable our team to refine the MMI system based on user feedback. Participants will be instructed to use the MMI system for several months so that the MMI system can learn about the participants' natural medication taking behaviors. Based on data collected during this period, our team will then develop and clinically validate specific intervention approaches targeting reasons for non-adherence.

The MMI system offers both flexibility and personalization in delivery of interventions. For example, modules may include delivery of brief content (e.g., text, audio, video) through the mobile phone via an application or through the smartwatch, phone calls, and text messages providing personalized content, or contact with healthcare providers or other significant people in the patient's life. Modules will be delivered at the precise time that the intervention is needed. For example, patients who are at risk for not taking their medication at the time of a side effect, such as severe joint pain, will be provided with one or more intervention modules that addresses that issue when the experience of pain is detected by the MMI system (e.g., educational content for addressing joint pain).

We will examine the relationship between the intervention and the changing medication-taking behaviors (as measured by MEMS devices) under different environmental,

personal, and behavioral contexts. Once the MMI system is evaluated with respect to feasibility, usability, and efficacy, it should be evaluated in a larger trial.

Chapter 3

Case Study: Understanding Patient Attitudes Toward mHealth Devices for Symptom and Adherence Tracking

3.1 Introduction

In recent years, new treatments have become available which have improved survival rates in lung cancer patients. One promising treatment option is the rapidly growing field of oral targeted therapies, which employs drugs that interfere with specific molecules involved in the growth, progression, and spread of cancer. However, these therapies can cause a variety of symptoms and adverse events that can impair quality of life. mHealth technologies may help individuals with lung cancer better track their side effects and manage medications on a day-to-day basis. However, understanding patients' attitudes toward smart devices such as smartphones, smartwatches, and

smart pill bottles, as well as their specific needs when using these devices, is critical before design and deployment studies of medication adherence can be carried out.

Contributions

In this chapter, we present our interview study with 9 individuals with stage III-IV lung cancer at an National Cancer Institute-designated comprehensive cancer center in the Mid-Atlantic region of the United States to assess the feasibility of using such devices for managing medication and medication related side-effects. We evaluated patients' attitudes towards the design and function of smart devices and how these devices fit into their daily life. Our results may help clinicians and researchers to co-develop effective mHealth system deployments for side effect and medication management in oncology populations.

This work is based on our study currently under review in Pervasive Health as of September 13, 2022 [105].

3.2 Background and Motivation

Lung cancer is the second most common type of cancer and the leading cause of cancer death worldwide, with over 2 million cases newly diagnosed each year [106]. The significant impact of lung cancer on the global population has led to the development of new targeted oral anticancer medications, which patients tend to prefer for their convenience over intravenous chemotherapy [107, 108]. While promising for survival

outcomes, these new therapies are commonly associated with adverse events (AEs) such as rashes or edema. AEs can lead to worsening symptoms, dose reductions, and even medication discontinuation if left undetected or untreated [109]. Researchers and clinicians are increasingly seeking more accurate ways to track medication-taking, monitor side effects, and detect possible AEs among patients taking oral anti-cancer medications at home, such as individuals with lung cancer.

Devices, such as smartphones, wearable sensors (e.g., smartwatches) and medication event monitoring systems (MEMS), enable direct, unobtrusive collection of clinically relevant behaviors *in-situ*. Mobile health (mHealth) and human-computer interaction (HCI) studies have shown that these "smart" devices are less prone to errors than traditional self-reports [109] and have established the usefulness of smart devices for medication and symptom tracking in daily life [110, 111]. At the intersection of HCI, mhealth, and oncology, smart devices have been shown to encourage medication adherence to oral chemotherapy [112], help patients feel more in control and informed about their care [113, 114, 115], and help clinicians feel better able to monitor patients' symptoms and tailor treatment accordingly [115]. While the majority of studies in mHealth and oncology have focused on physical activity tracking for breast cancer patients [116, 117], several recent studies have focused exclusively on medication and symptom tracking for lung cancer. LuCApp is a mobile application for patients with lung cancer that sends automated reminders to complete symptom logs as well as questionnaires related to quality of life and support needs [118]. A proposed randomized controlled trial for the app will examine the impact of side-effect tracking on quality of life. A randomized controlled trial has also been proposed for

SYMPRO-Lung, a web application for lung cancer patients that leverages patient-reported outcomes (PROs) for symptom monitoring [119].

Despite the promise of smart devices for symptom and medication management, many challenges remain unaddressed. A 2017 study of medication adherence technologies such as smartphone apps among older adults showed that adherence was impacted both by participants' schedules and the symptoms they experienced [111]. Further, a study of medication tracking among patients with atrial fibrillation uncovered issues such as the inability of smart pill bottles bottles to integrate into patients' existing routines [73]. These challenges highlight the importance of further investigation into patients' perceptions of smart device use during treatment, so that future interventions designed to improve adherence can be customized to patients' individual needs. Further, Social Cognitive Theory can help organize patients' needs into unique constructs - personal, environmental, and behavioral - so that interventions can be targeted for specific construct(s). For instance, an app might be designed to provide additional social support, thus targeting *environmental* factors within the context of cancer treatment.

In this work, we address the following research question ***(RQ): What attitudes do patients with lung cancer have toward smart device use for managing their medications and tracking their symptoms?*** We present the results of a cross-sectional, qualitative study in which we conducted semi-structured usability interviews with 9 individuals with stage III-IV lung cancer receiving treatment at a large university cancer center in the Mid-Alantic region of the United States. Our results give insight into patients' preferences and priorities regarding the use of smart

devices as part of their self-management routines during cancer treatment

3.3 Methods

This study was approved by the Institutional Review Board for Health Sciences Research (IRB-HSR) at the University of Virginia, and the study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice standards. Patients provided written informed consent prior to enrollment and participation.

3.3.1 Recruitment

Using purposive sampling, we recruited patients at an National Cancer Institute (NCI)-designated comprehensive cancer center in the Southeastern United States. All patients were 18 years of age or older and were being treated for advanced stage non-small cell lung cancer (NSCLC) with EGFR mutations or ALK gene rearrangements and were receiving oral targeted therapies (tyrosine kinase inhibitors [TKIs]) as part of their treatment. Patients were first identified for inclusion by the sixth and seventh authors, who are practicing oncologists. The first author attempted to contact prospective participants both in clinic and via telephone calls, and provided interested individuals with a secure, electronic consent form to sign. We approached 40 patients in total, 23 of whom either explicitly declined prescreening or were unreachable after one or more attempts to contact them. 17 patients agreed to prescreening, and 11 ultimately consented to participate in the interview study.

Two participants did not respond to study coordinators' efforts to schedule the study interviews after consenting, bringing the final number of participants to 9.

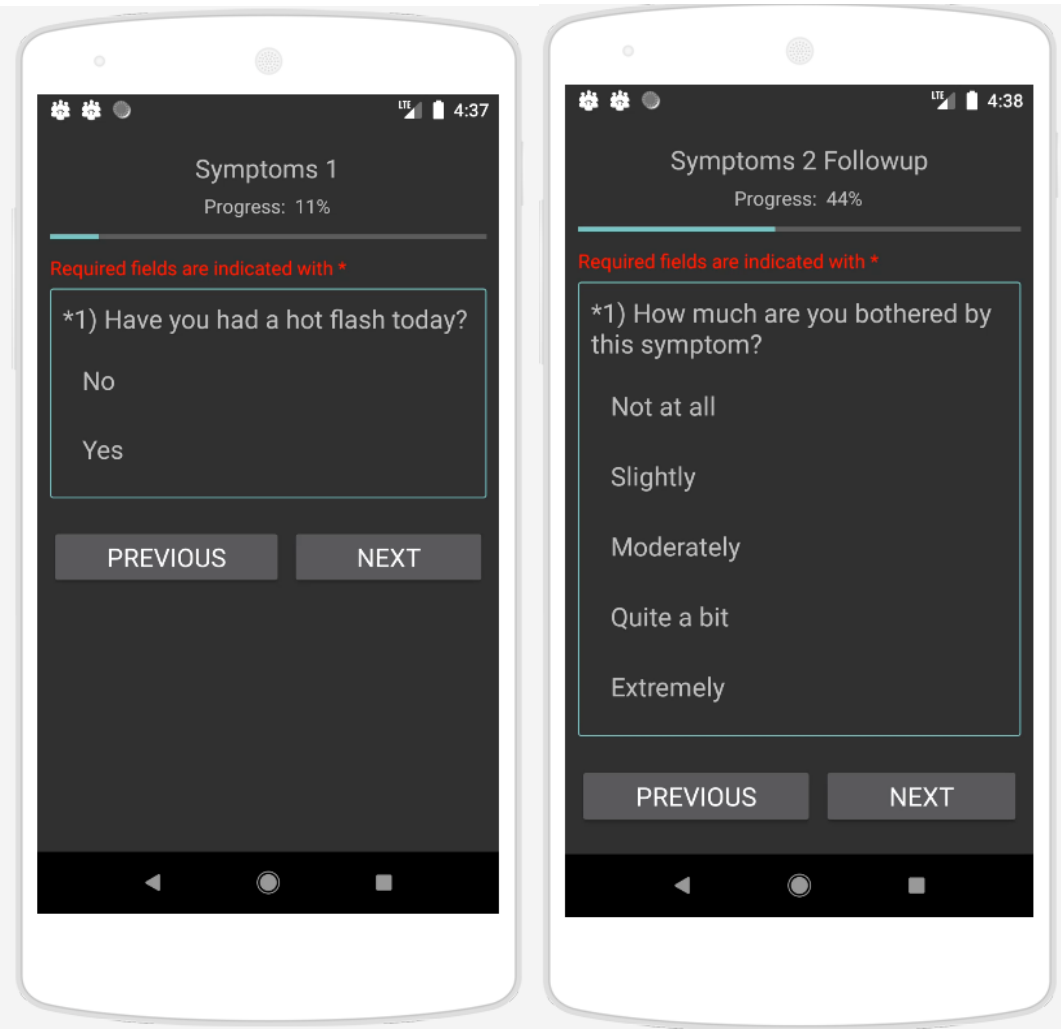
3.3.2 Data Collection

Using a standardized interview guide, we conducted semi-structured interviews with 9 participants between September 2020 and July 2021. Out of an abundance of caution during the COVID-19 pandemic, interviews were conducted by one interviewer remotely via a HIPPA-compliant version of Webex [1]. We administered a secure, online demographics survey via Qualtrics [2] at the end of each interview. The first interview (with P1) focused on smartphone use and an interactive demonstration (demo) of a smartphone app emulator. P1's interview informed the design of subsequent interviews, which included an interactive demo of a smartwatch app emulator and a researcher-guided demo of a smart pill bottle cap in addition to the original smartphone app demo. In this section, we describe our process for each device demo in detail. The full list of interview questions is included in Appendix A

Smartphone

We created a high-fidelity prototype of Sensus [32], a smartphone application that gathers passively sensed indicators of human health and behavior (e.g., location, heart rate, and skin temperature). In deployment studies, Sensus can also be used to

[1] https://www.webex.com/
[2] https://www.qualtrics.com/



(a) (b)

Figure 3.1: Sensus Prototype showing EMA survey about symptoms.

gather real-time participant feedback using ecological momentary assessment (EMA), a method of gathering data in which participants are polled in real time in order to avoid recall bias [22]. EMAs are commonly delivered via digital methods such as text messages or push notifications from mobile apps. We created a Sensus study

protype with EMA surveys about participants' quality of life (e.g., sleep, symptoms, and side effects) and day-to-day activities (e.g., location and socialization). We then loaded the Sensus protocol via Appetize.io [3], an app demo platform for the web; the prototype is shown in Figure 3.1. During the interview, we used screen sharing to show participants how to scroll and select questions within the prototype; this was necessary, as using a mouse for these tasks is very different from swiping and tapping with one's finger on a real smartphone. We then asked participants to use the prototype via screen sharing to practice answering the survey questions. Finally, we asked participants 15 follow-up questions that covered their perceptions about the experience of filling out surveys on a smartphone, the relevance of the smartphone to their current medication management routine, and their willingness to use Sensus on a smartphone for a long period of time.

Smartwatch

We used a web-based prototype of a smartwatch application (shown in Figure 2) that was preloaded with quality-of-life questions, as well as questions regarding activities (e.g., "Are you exercising right now?"). We explained to participants how the smartphone could be used to send EMAs and notifications to the smartwatch. We then asked participants to use the smartwatch app prototype via screen sharing to practice filling out EMAs. Finally, we asked participants 15 follow-up questions that covered their perceptions about the experience of filling out EMAs on a smartwatch, the relevance of the smartwatch to their current medication-taking routine, and their

[3] https://appetize.io/

(a) (b)

Figure 3.2: Smartwatch app prototype showing EMA survey about symptoms.

willingness to use the app on a smartwatch for a long period of time.

Smart MEMS Cap

For the final segment of the interview, we used screen sharing to demonstrate the use of the RxCap [4], a bluetooth-enabled MEMS cap that records each time the cap is unscrewed as a medication-taking event. Since most participants had not seen or used a smart pill bottle before, we first explained the purpose and function of the

[4]https://rxcap.com/

cap. We then showed participants how one would remove the cap to take medication, and how the cap would blink and beep upon removal. We also explained how the cap could be connected to an application on the user's smartphone to help keep track of when they took their medication. In order to gain a better understanding of participants' medication-related needs, we asked questions about the types of medications participants were currently taking, the frequency with which they took them, their preferred storage method (e.g., pill box; original bottle) and location (e.g., in the bathroom; on a nightstand), and what kinds of alerts or reminders they used to help them remember to take their medications (e.g., app; phone alarm) We then asked 15 follow-up questions about the relevance of the cap to their current medication management routine and their willingness to use a smart pill bottle to store their medications for a long period of time.

3.3.3 Data Analysis

Interviews lasted between 30 minutes and 1.5 hours and were audio recorded. The first and second author transcribed the interviews verbatim. The first author then applied initial codes from the interview transcripts to develop a preliminary codebook. She then worked with the second, third, and fourth authors use an iterative, inductive approach to refine the initial codes, develop new codes, and extract the overall themes. Iterations continued until all coders reached consensus.

Demographic data were gathered and analyzed by the second author, who extracted the data from the secure Qualtrics survey and aggregated it. Aggregation was done

at the population level to protect participants' privacy.

3.4 Results

3.4.1 Participant Demographics

Participants' ages ranged from 33 to 86 years with a gender distribution of 6:3 (female: male). Most participants self-identified as White (8/9, 89%), followed by Asian (1/9, 11%), and Hispanic (1/9, 11%). All participants had been diagnosed with lung cancer at least 6 months prior to the study. Among those who reported their lung cancer stage, the majority were diagnosed with stage IV (7/9, 78%), followed by stage III (1/9, 11%). Less than half (3/9; 33%) were former smokers. Among the former smokers, the average number of years of tobacco use was 22 years.

3.4.2 Interview Study Findings

Participants' attitudes, concerns, and needs regarding smart device use spanned four key thematic areas: *device and application design*, *lifestyle*, *abilities*, and *obligations*. In this section, we delve into each of these themes in detail to answer our original research question **(RQ)**: *What attitudes do patients with lung cancer have toward smart device use for managing their medications and tracking their symptoms?* In this section, we present participants' attitudes and concerns toward the individual devices as well as their needs and preferences for notifications they might receive from

any device. We also describe the how participants' personal and social obligations factor into their willingness to use smart devices. Finally, we highlight the roles of self-efficacy and obligation to self and others in motivating smart device use.

Device and Application Design

Smartphone and Smartwatch. Participants appreciated the smartphone and smartwatch for their compact size and ease of use. Both P1 and P11, for instance, expressed a preference for the smartphone over bulkier technologies. In P11's words, "*It doesn't force me to have to go to the computer.*" While the smartphone was familiar to most participants, the smartwatch was not, and participant opinions on the smartwatch were divided. Some were drawn to the smartwatch because they could input data directly on their wrist in a discreet way. P4, for instance, described how his personal smartwatch was useful for discreetly checking messages while at work. P11, however, worried that smartwatch notifications in particular were a privacy risk. Given that the watch must be worn at all times rather than kept aside in a purse or pocket, the watch has the potential to draw attention to private messages in social settings:

> "*Notifications will bother me more on my watch than on my phone…I'm not sure if I would like going out for dinner, and all of a sudden noticing that my wrist is lighting up with a message and somebody across the table says, oh, you have a message in your wrist… [Or] what if I'm having an important meeting with somebody?*"

P4 also found the smartwatch somewhat intrusive (despite regularly using his own), and preferred to keep the device in the background as much as possible: "*It requires an answer right then and there. Personally I don't like inputting [data]. I see [the watch] as more of a way to receive information.*"

Participants also mentioned several design-related needs and concerns, with regard to the smartwatch and smartphone. Some worried the surveys were too long and would become cumbersome, or that the device's battery might drain too quickly due to running an app. Attitudes about device size were divided; one participant found the smartwatch screen too small and hard to navigate, while another, P6, found the watch to be too big for regular use: "*I'm not a big fan of wearing much on my wrist..I would be inclined to forget wearing it, I'm afraid, because it's bulky… I do not like intrusive technology.*"

Several participants wanted changes to the surveys, including more aesthetically-pleasing color schemes and the ability to comment on the frequency of their symptoms. Participants also wanted to complete surveys at their own time and pace, with several wanting to set their own notification schedule.

Pill Bottle. Of the three smart devices we studied, the smart pill bottle received the least support from participants. Most participants were taking multiple prescription medications in addition to multiple vitamins and supplements, and disliked the smart pill bottle because it could only hold one type of pill at a time. As P11 described, they tended to prefer divided pill boxes for everyday use: "*I think the box that has the separated days would be more useful because it would be recording not only that*

you took it, but [when] you took it…So the data that you would record would be more complete."

Notifications. In general, participants valued their privacy and peace. They wanted notifications to be unobtrusive and discreet, especially in public settings such as the workplace. Participants' preferences, in this regard, were very personal. Some participants preferred to leave most notifications off and found them "annoying". P8 was willing to receive vibrations only, in keeping with his work obligations: "*A vibration is best, because a lot of times I'm in management meetings. Obviously, we all have our phones turned down.*" Others, such as P7, were willing to receive audible "dings" on any device, provided they were not overly loud or repetitive:

> "*I would want it to be quieter and more subtle, …and also not persistent. So one notification is fine, [but] five notifications would not be fine… I would not want to have to keep seeing it. I'm [also] notorious for clearing out my notifications, and in fact I turn off a lot of my notifications because it's a privacy issue to me.*"

P6 expressed a similar preference: "*I set my alarms for my meds so I would definitely [want notifications]…I would probably have it be a single ding…I don't want anything irritating like 'DA, DA, DA, DA!' Just like a single ding would work.*"

Several participants also expressed a desire for survey notifications to be integrated with their electronic health record apps, so that all their health-related notifications showed up in one place. For example, P4 described how he would be more likely to take surveys when checking for appointments in MyChart, a popular electronic

medical record (EMR) application [5]. Importantly, participants wanted to receive notifications only when absolutely needed, given the burden of time their cancer treatment already placed on them. For instance, P11 described how her view of time had changed since her diagnosis and re-emphasized that notifications should be minimally disruptive, especially during family and personal time:

> "*When you have cancer, too many things seem too trivial, and you want to concentrate every day on using your time in the best possible manner. So it's funny, I'm now quite bothered by all these notifications that come to me about celebrities…but I do want to get notifications if my sons do something. So I think it depends. I would say not too many; enough notifications that we can do this [study], but not unnecessary notifications.*"

Lifestyle

Participants' lifestyles heavily influenced their attitudes towards smart device use. P6, echoing many other participants, cited her familiarity and current use of smartphones as a reason they would be willing to use the Sensus app as part of a future study: "*Like a lot of people, I use my phone more than any other device.*" Participants were less familiar with the smartwatches. Even P4, who owned a smart-watch, was concerned about learning to use a different type of smartwatch when he already owned one that worked well for him:

[5] https://www.mychart.com/

> "*I wouldn't like it. I prefer my watch and the features my watch has, I've already gotten used to it. I wouldn't want to learn a whole new system, and I'm assuming if it's a research watch I wouldn't be able to install any of my own apps on the watch anyways.*"

Participants' schedules and levels of flexibility varied, though most concluded that they were more available on weekdays than on weekends. Weekends were often reserved for family time and other social activities (e.g., hosting friends or attending church services). For instance, P1 described the importance of time with her husband: "*We're out hiking or doing projects. I'm less likely to think about doing something like a survey.*" Similarly, P6 did not want to be interrupted during her valued social time:

> "*Generally we are busier on the weekends with catching up with friends and family…I set an alarm on my phone for taking my medication, and then honestly if we're out socializing, I end up snoozing the alarm and snoozing the alarm… I take it within an hour or two. But… [at] nine o'clock on Saturday night I don't generally want to be interrupted with a reminder about something, and I think I'll feel the same way about the surveys.*"

Several participants also mentioned that their activities could put them out of range for receiving push notifications from a study device (e.g., alarms or reminders to take a medication dose). For instance, before the COVID-19 pandemic forced many people to stay at home, P8 liked to go hiking on the week-ends in remote areas

with little-to-no cellular service: "*Right now we're all hunkered down [during the pandemic]. I used to be out of range of any technology if I was skiing or backpacking. I could be gone for 13 days, out of range.*"

Throughout the interviews, participants repeated their commitment to habits and routines as a major influencing factor on their attitudes toward smart devices. Several participants, such as P6, assured us that using the devices would "just become a part of a daily routine". P9 even likened smart device use to taking medicine regularly: "*What's the difference of using the app every day and then using the medicine every day? I don't know if there would be a difference.*" While participants felt the smartwatch and smartphone could fit into their existing routines, they did not feel the same way about the pill bottle. The reasons participants gave for not wanting to use the pill bottle were as much of a lifestyle concern as a design concern. Namely, managing multiple pills had driven participants to establish longstanding, personalized medication-management routines that already worked well for them. P2 put it simply: "*I don't need [the smart pill bottle] – "I have no problem with what I'm doing now.*"

Abilities

Participants exhibited varying degrees of self-efficacy with regard to using smart devices as part of their medication management routine. P2 noted that, while she was confident she could use the real, physical devices, the screen size and difficult scrolling in the online prototype were challenging for her. Others such as P1 expressed confidence in their own technical skills, but were doubtful that others, especially in

older age groups, would be able to use the devices: "*I don't think I would have any problem at all using it. I could see if it was my mother, I would have to train her how to use it. But the way it's set up, for anyone that uses apps, it's pretty obvious what to do.*" Still others believed they would need significant support from the study team or another support person. For instance, P7 asked, "Are you going to train me really well in how to use that smartwatch?" P3 believed she could use the smartphone if she received outside help from her grandchildren or a tutor: "*I might even hire a tutor to help me…I'm not that great on a smartphone for sure.*"

Obligations

Besides force of habit, participants cited obligations to their doctors and to themselves as a factor influencing their willingness to use devices. Participants took great pride and responsibility in managing their health as best they could during treatment. They valued smart devices for their ability to track symptoms over time, and wanted the ability to share this information with their providers. For instance, P1 described how an app could help her keep an accurate record of her symptoms to present to her oncologist: "*I'm not going to call my doctor and say, oh I had a mouth sore today. But if I had an app, if I was supposed to use it every day, I would put it into the app.*" Likewise, P7 would be willing to fill out surveys more frequently if it helped with symptom management: "*If [there were] questions that I felt like were important for me and my doctor to know, I would do it five times a day… I would do it… if I thought it would help manage my symptoms.*"

3.5 Discussion

Our study highlights a number of considerations and challenges for designers at the intersection of mHealth, medication adherence, and oncology. Participants' openness to using certain devices may be mitigated by *personal* factors such as their familiarity with the device and even their fashion preferences, by *behavioral* factors such as confidence in their technical abilities, and by *environmental* factors such as work and social obligations and whether they will be available and within range for receiving push notifications. Based on these findings, we present several practical design considerations in the following sections.

3.5.1 Give participants agency over their notifications

Notifications and reminders to take one's medication play an important role in interventions for individuals with cancer. Prior work has demonstrated the feasibility of using information from smart devices to inform the delivery of missed dose messages via EHRs such as MyChart [73]. This approach is a promising step towards more personalized notifications and interventions. Prior research with cancer patients taking oral chemotherapy has emphasized the importance of taking the user's schedule into account when delivering reminders [120]. In a similar vein, participants in our study overwhelmingly expressed a desire to control the timing and format of notifications as a condition of using the smart devices in their regular medication-taking routine. We recommend that designers of smart device applications for medication management give participants a range of options for customizing their notification

frequency from within the app. Designers might consider setting a default schedule based on times of day most commonly associated with medication taking and alertness, then enabling participants to customize this schedule as needed. For instance, 8 AM and 8 PM often correspond with morning mealtimes and evening bedtime routines, respectively. Ideally, users would be presented with a screen that allows them to set the exact day(s) and time(s) they would like to receive notifications, as well as the type of notification (e.g., vibration, beep, or banner) for each day and time. Giving participants agency over their notifications in this manner is a small cost for designers, but a major step toward protecting participant privacy and ensuring notifications are well-integrated into participants' daily routines (rather than intruding on them).

3.5.2 Advocate for better avenues for secure sharing of patient-entered data with clinical care providers

Prior works have established patients' openness to sharing information such treatment satisfaction and adverse effects with their clinicians via apps, provided that the information can be used to complement their treatment [121]. Our participants shared this attitude. Given the significant physical and emotional toll of their lung cancer treatment, participants had a vested interest in being able to review and share as much of their passively- and actively-sensed health data as possible with their doc-tors. Researchers have called for better infrastructure for secure information sharing between between patients and providers in the context of mHealth for

cancer care, given the limitations of current modalities [122]. Yet, this remains an open challenge. Electronic Health Records (EHRs) are the gold standard for health information and records management in our digital world, yet they are primarily designed for displaying information entered by clinical care providers (e.g., patient lab values and test results) and for facilitating basic secure messaging between patients and providers. EHRs in their current form are not equipped to receive and process information from consumer devices such as smartwatches or from custom medication and symptom tracking apps, in part due to strict requirements imposed by laws such as the Health Insurance Portability and Accountability Act (HIPAA; [123]) and the Health Information Technology for Economic and Clinical Health Act (HITECH Act; [124]). Moreover, building a standalone application that securely transmits patient-entered data to the patient's healthcare provider via an existing EHR platform be an enormous challenge. Such an application would not only need to comply with current market standards for secure health information sharing, such as Health Level 7 (HL7) [125], but would require direct collaboration with leading EHR vendors. Additionally, such an application would need to navigate the gaps left by HIPAA with regard to digital healthcare tools [126].

Designers seeking a short-term solution for patients who wish to share information with their doctors could provide in-app visualizations at different timescales, such as daily charts of the times of day a patient took their medication, or weekly and monthly graphs of adherence percentages over time. These should be easily-exportable to images that participants could save to their smartphone and share with their doctors manually in-clinic during routine appointments. Designers taking

this approach should consult with both patients and clinicians when designing such visualizations, to ensure they are clear and concise for both parties. As a long-term solution, designers should consider advocating for improved guidance on patient-to-clinician information sharing via digital technologies, at the national level, and should seek out long-term collaborations with EHR vendors and smart device manufacturers where possible.

3.5.3 Ensure participants have access to adequate support resources during deployment studies

Prior work in mHealth has shown that patients with lung cancer may feel they are lacking sufficient support and self management skills, in regards to their disease [127]. Several participants echoed these concerns specifically with regard to using smart devices. Indeed, many expressed a hesitancy to use smart devices due to their perceived lack of technical proficiency. Participants also expressed a need for extensive support from the study team, should they choose to use the smart devices in a future deployment study. To increase participants' confidence, we suggest providing a comprehensive technology use manual and other written educational materials that describe how to use each study device. We also recommend that study team members review these materials with participants, and provide in-depth demonstrations of each each device. We note that technological concerns are likely to arise during deployment. To adequately address these concerns, we also recommend providing the participant with a specific study contact designated to addressing and supporting

individual technology needs.

3.6 Limitations

This study is novel given its focus in evaluating perceptions of smart technology among individuals living with advanced lung cancer. However, it does have limitations, including the small sample size. We faced several recruitment challenges during this study. Cold-calling potential participants was largely unsuccessful. Among those who responded to cold calls or were willing to speak to us in-clinic, most declined. Their reasons included disease burden (e.g., fatigue from treatment), busyness due to participating in other research studies, and lacking a computer for the study interview. Whether these challenges were unique to our study population is outside the scope of this paper; however, we recommend that future studies cast a broad recruitment net across multiple treatment facilities if possible.

Like most studies conducted during the COVID-19 pandemic, we also faced challenges in adapting our study activities to be fully remote. While we were able to con-duct recruitment in-clinic by taking many precautions such as masking, we opted to conduct the interviews remotely to reduce the risk of transmission to participants. This decision fundamentally altered who we enroll in the study. Our remotely conducted interviews required a personal computer (PC) with a mouse and microphone, preventing those without a PC from participating. Moreover, those who did participate did not get the in-person experience of physically interacting with the study devices. Additionally, we struggled to recruit a racially diverse sample. We also urge

researchers in the field to continue to increase efforts to recruit diverse participants. Given the significant racial and social disparities present in the incidence of lung cancer [128] and other diseases, diverse samples are necessary for designing mHealth tools that serve as many patients as possible.

Chapter 4

COMP-SCT: A Computational Framework for Social Cognitive Theory

4.1 Introduction

The wealth of information contained in even modest amounts of passively-sensed data makes it possible, now more than ever, to construct robust digital biomarkers of the patient experience, and computational approaches such as machine learning hold the key to unlocking the patterns within this data. The MMI system, presented in Chapter 2, lays the groundwork for tying raw sensor data to patient experiences via passive sensing, and for developing and deploying personalized mobile interventions for medication and symptom tracking. Further, our interview study presented in Chapter 3 helps hone in on the specific personal, behavioral, and environmental factors that may impact adherence to mobile interventions, so that these interventions can be customized for specific patient needs. Importantly, however, we still lack

a comprehensive process for computing relevant features (with regard to medication adherence) from raw sensor data.

While many generalizable computational frameworks for digital phenotyping exist, few focus exclusively on SCT. Moreover, the few computational frameworks that do focus on SCT are concentrated in dynamical systems and control systems theory [129, 130] and tend to focus on the mathematical relationship between SCT constructs. There exists a major opportunity to develop a new framework that explicitly guides the translation of raw, passively-sensed data into features for predictive analysis.

In this chapter, we present COMP-SCT, a framework for deriving personal, behavioral, and environmental features of user context at multiple time scales using Social Cognitive Theory. COMP-SCT draws inspiration from leading digital phenotyping frameworks [5, 131, 132, 133] and the MMI system in order to map SCT constructs to relevant temporal features for predictive analytics. Importantly, COMP-SCT's novelty lies in its explicit grounding in Social Cognitive Theory which, to date, has been largely disregarded in the digital phenotyping literature.

In the following sections, we describe COMP-SCT in detail and apply it to two case studies of breast cancer patient data. The first case study focuses on behavioral and personal features (i.e., user engagement and mood, respectively), while the second focuses solely on behavioral features (i.e., use of a smart pill bottle cap).

Content in Sections 4.2 and 4.3 has been adapted from our study of engagement and mood among breast cancer patients [134]. Content in Section 4.4 is forthcoming

work.

4.2 The COMP-SCT Framework

COMP-SCT's steps are outlined in Figure 4.1. Our process is informed by best practices in machine learning (and, more broadly, data science) for extracting and analyzing features from raw, multimodal data.

4.2.1 Preprocessing

Invalid Data Removal

Preprocessing is critical for preparing the raw data for analysis and typically begins with removing invalid data. Invalid and missing data are common to all mHealth datasets and can occur due to user error, sensor malfunction, or lack of user action. This may be particularly relevant in the context of breast cancer patients, given the demands and cognitive effects of treatment (e.g., chemotherapy). For example, a GPS sensor may provide an inaccurate reading, or a user may complete an ecological momentary assessment on their phone but fail to click the "submit" button. Large swaths of invalid or missing data can degrade the quality of the dataset and lead to less accurate analysis, making it imperative that researchers handle both with care. In mHealth studies, invalid data is best described as data that falls outside the acceptable range for a given variable. One example is app launches that are too

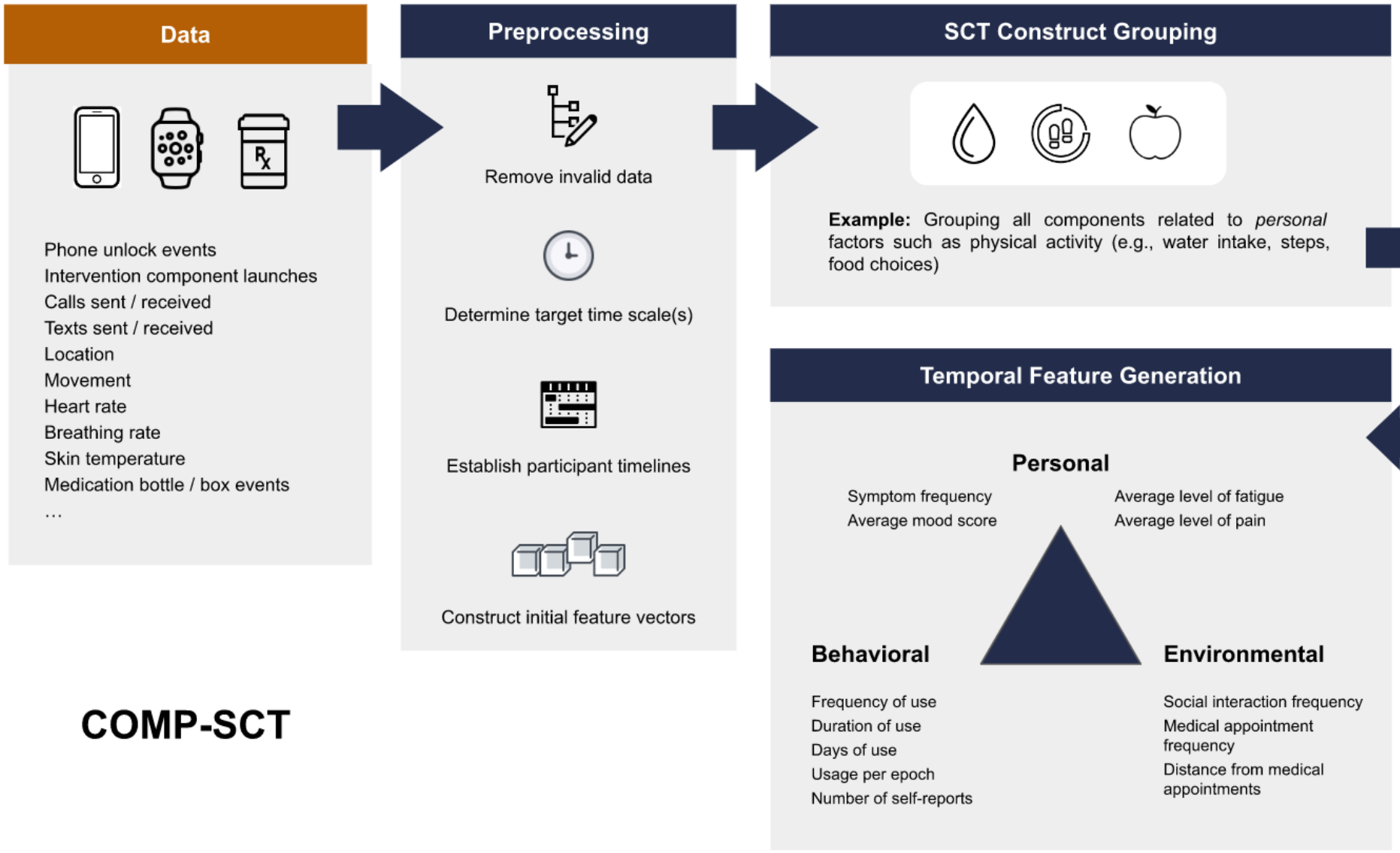


Figure 4.1: COMP-SCT Framework

short (e.g., less than 5 seconds) or too long (e.g., greater than 5 hours) in duration. In the former case, the user likely opened the application and immediately closed it, and thus did not use the app; in the latter case, the low-level mobile operating system code that monitors app usage may have failed to record the end of the user's usage activity period for the given app. Invalid data should be removed at the very beginning of the preprocessing stage in order to reduce the complexity of the dataset and the computing power needed to analyze it.

As preprocessing steps are certainly not unique to mHealth datasets, we refer the reader to García et al. [135] and García et al. [136] for further reading.

Feature Vector Construction

mHealth studies often leverage data from multiple sources (e.g., GPS, Accelerometer, EMA), all of which may have different sampling rates. In order to formulate a meaningful prediction task, researchers must transform the raw, multimodal sensor data into standardized vectors of data (*features*) for specific units of time (e.g., one vector per week). The unit of time is dependent on two things: The target variable (*outcome*) we want to predict, and the time scale(s) along which we expect this outcome to change. Consider a study in which we want to predict whether an individual with breast cancer is anxious at some future point in time, and we want to do so using features of engagement from a mobile phone. We can reasonably expect mood to fluctuate on a daily basis, based on prior literature in mood prediction []. Moreover, we might also want to predict mood for less granular time periods (e.g., weeks) in order to uncover overarching patterns in the data. Thus, we may choose to construct vectors at both the daily and weekly time scale.

Importantly, vectors should be aligned to a standardized timeline based on each individual's start date (i.e., the day on which they started using the intervention).

Researchers often face challenges in recruiting cancer patients to enroll in trials of digital interventions [137], and thus rely on a rolling enrollment period to increase recruitment over time. As a result, mobile health datasets collected from cancer patients often have different coverage periods for each patient. Start dates must therefore be converted into relative timestamps, with respect to the study length, in order to establish a standardized timeline for analysis. Consider two participants, Participant A and Participant B. Participant A begins the study on 01/01/21 and submits a self-report via a mobile application on 01/02/21; Participant B begins the study later, on 01/15/21, and submits a self-report on 01/20/21. Despite their different start and submission dates, both participants are said to have submitted data during the first week of the study. This is just one example of how timestamps may be aligned, as researchers may wish to use a different temporal granularity (e.g., day of study).

4.2.2 SCT Construct Grouping

Observational and intervention studies of user context may use an app or suite of apps to gather user data, and each app may have several distinct modules. In this case, the apps and individual modules are all *components* of the intervention. Researchers must decide whether to analyze intervention components in aggregate, individually, or for SCT construct groupings of apps or modules. Increasingly, researchers are developing suites of related apps that all target a general domain of health, such as mental health, but which have distinct target goals. In the IntelliCare suite [131], for instance,

the Thought Challenger app helps users address negative thoughts, while the Daily Feats app helps users track their accomplishments and stay motivated. Patients with diseases such as breast cancer may benefit from multiple apps or a suite of apps given their unique physical, emotional, and social needs tied to their disease. Multiple apps (or modules within a single app) that serve these different needs independently may be necessary for providing adequate support during treatment. Thus, it may make sense to group components by their domain and map this domain to an SCT construct. For instance, mental health components of an intervention would fall under the *personal* factors construct. Custom groupings may also be beneficial. For instance, components could be grouped according to a cutoff score for a metric such as usage frequency (e.g., "highly-used apps" are a group containing all apps used 6 or more days per week).

4.2.3 Temporal Feature Engineering

Traditionally, researchers have measured SCT constructs such as engagement using blunt usage metrics such as the total number of app sessions over the course of an intervention, or the number of users that fail to complete an intervention [16]. However, with the increasing ubiquity of sensor-equipped smart devices, researchers have been able to derive more temporally granular features of user context from logs of phone or app usage [138]. Time segmentation of behavioral features has been used to detect human activity and behavioral patterns broadly, breathing state changes [139], social behavior [9, 140], loneliness [5] and sleep disruption events [141].

Previous works within mHealth, specifically, have used theory-driven segmentation to examine context at hourly intervals, across multi-hour spans (e.g., "morning," spanning 6 am to 11:59 am), and at weekly intervals [5, 9, 11, 44]. Several common behavioral features have emerged from these and related studies, including frequency of use of intervention components (e.g., how many times per week a user opens an app), number of days of use, duration of use, whether any use occurred in a given time period (*epoch*; [27]), and number of self reports submitted [5, 44, 51, 142]. For a summary of these and other "analytic indicators of engagement," we refer the reader to Pham et al. [133].

When constructing temporal features, researchers should weigh the nature of the condition being studied and, in turn, the time scale(s) along which SCT constructs are likely to vary. Women newly diagnosed with breast cancer may have only sporadic pockets of time throughout the day to engage with a mental health app due to increased time spent attending doctor's appointments and managing their illness and sequelae of related factors. Additionally, due to the disruptive impact of anxiety, depression, and cancer treatment on daily rhythms [143], breast cancer patients experiencing mental health challenges may engage with mental health apps at irregular times. Given the stressors that breast cancer patients face, constructing vectors for short and frequent units of time (e.g., days, or times windows in a given day) may be most appropriate in order to capture fluctuations in mood.

Below, we modify table 4.1 to map exemplar SCT-related features to suggested minimum time scales for measuring these data, with a focus on features relevant to breast cancer studies. We also include justifications for these suggested time scales

Table 4.1: SCT Constructs, Measurement Methodology and Exemplar Features

SCT Construct	Data	Features	Min. Time Scale(s)	Literature
Personal	Hot flashes, stress	HRV, body temperature, breathing rate, EMA	Continuous	[144, 145, 146, 147]
	Mood, side effects (e.g., rash), fatigue, pain	EMA	Days	[148]
Behavioral	Eating, medication-taking, exercising (e.g. walking or running), sleeping	HRV, movement velocity, movement variation, location	Continuous	[82, 84, 87, 149]
Environmental	Engaging with a mobile intervention (e.g., using an app)	App use frequency and duration	Multi-hour spans (e.g., "morning")	[44]
	Social interactions - Non-medical (e.g., meeting friends)	Audio signals, social media activity, semantic location diversity, text and call frequency, app use frequency	Continuous	[150]
	Social interactions - Medical (e.g., speaking with doctor)		Days	[73]

from related studies from mHealth and oncology:

We urge researchers to balance the need for temporal granularity against their dataset size. Larger datasets with more frequent measurements will naturally allow for more time windows (e.g., hourly). Researchers should also take care to ensure epochs are neither too broad nor too narrow. Epochs that are too broad will fail to capture meaningful patterns, while epochs that are too narrow will introduce sparsity into the dataset and decrease the effectiveness of the analysis.

4.3 Case Study: Predicting Mood from Passively-Sensed Features of User Engagement

4.3.1 Background and Motivation

In the United States, 1 in 8 women will receive a breast cancer diagnosis at some point in her lifetime [151]. Breast cancer is currently the leading cause of cancer death in women [152]. Patients with breast cancer encounter a range of psychosocial stressors that extend beyond the physical effects of anticancer treatment, including emotional distress, diminished well-being, and increased symptoms of depression and anxiety [153, 154]. Untreated symptoms of depression and anxiety in women with breast cancer can lead to poor quality of life [155], increased mortality [156], and high economic costs [157].

Interventions that emphasize skill acquisition, such as cognitive behavioral therapy, have been shown to effectively reduce symptoms of depression and anxiety in patients with breast cancer [158, 159]. However, numerous barriers prevent patients with can-

cer from receiving adequate treatment, including high financial [160] and time [161] costs, social stigma [162], and a severe shortage of trained psychotherapists, particularly in rural and underserved areas [163]. Combined, these barriers lead to almost half of breast cancer survivors reporting unmet psychosocial needs [164].

Increasingly, researchers are leveraging mobile phone apps to address mental health issues in patients with cancer. Apps are frequently cited as a way of extending cost-effective care [165, 166]. In many cases, digital interventions (i.e., web-based and app-delivered interventions) that mirror the content of in-person therapy perform just as well in reducing mood symptoms [167, 168]. App-delivered interventions can decrease barriers associated with traditional in-person interventions as treatment is affordable, is readily available, offers efficient use of time (i.e., no delays to begin treatment and self-pacing), and is no longer limited by factors such as geographic proximity to available psychotherapists. This is particularly relevant for women undergoing anticancer treatment regimens who may only have small pockets of unstructured time in a day. Numerous studies have validated the use of apps to reduce depression and anxiety symptoms [169, 170], including in patients with breast cancer.

Although access to high-quality treatment is a major issue that app-delivered interventions are well poised to address, sustained engagement is a common problem [16]. Engagement is critical as it is necessary for treatment success, as studies have documented a dose-response relationship in app interventions [171, 172]. A barrier to advancing knowledge of engagement in digital interventions is data density. It is common for app-delivered interventions to be deployed by a user when and where they are most convenient, potentially leading to a large data set. Fortunately, advances in

machine learning have made it possible to analyze vast volumes of engagement data. However, translating these raw engagement data into clinically meaningful observations is an ongoing challenge in oncology research using mobile health (mHealth) tools [116]. Moreover, to date, no studies have presented a clear process for analyzing the relationship between engagement with mental health apps and outcomes in cancer populations using machine learning.

4.3.2 Methods

To illustrate the app engagement process, data were extracted from a 7-week trial [173] of a mobile mental health app suite among women newly diagnosed with breast cancer (N=40 participants). IntelliCare [131] is a collection of apps that use an elemental, skills-based approach to improving mental health. In-app exercises are meant to be intuitive, requiring few instructions to complete, and most of these exercises can be found on the first screen presented by the app. Participants used their own personal phones and were recruited from a breast care clinic at a US National Cancer Institute–designated clinical cancer center. A detailed description of the recruitment method, as well as the goals of the IntelliCare apps, can be found in a paper that depicts the primary outcomes of the study[173]. Participants downloaded and tried 1 to 2 apps each week. All participants received light phone coaching that focused on addressing usability issues with the apps, which included an initial 30-minute call at the beginning of the trial, followed by a 10-minute call 3 weeks into the trial. Although 58% (23/40) of participants completed the intervention in the original trial,

because of technical issues exporting app use metrics from the system, detailed app engagement data were only available for 35% (14/40) of participants.

4.3.3 Ethics Approval

This study was approved by the institutional review board at the University of Virginia (UVA IRB-HSR#20403).

4.3.4 Participant Demographics

Participants had a mean age of 56.8 (SD 11.6) years; 82% (31/38) of participants who indicated their race were White, 11% (4/38) were Black, 3% (1/38) were Hispanic, 3% (1/38) were American Indian or Alaska Native, and 3% (1/38) were multiracial.
Measures

The Patient Health Questionnaire-4 (PHQ-4) [174] and Patient-Reported Outcomes Measurement Information System-29 (PROMIS-29) [175] were used to assess the symptoms of depression and anxiety at baseline and after the intervention. To allow for an examination of changes in mood symptoms over the course of the trial, a 2-item measure of symptoms of anxiety and depression was administered once daily during week 1 and at the beginning of weeks 2 to 6 of the trial. The daily measures from week 1 were averaged. This measure comprised questions from the PHQ-4 ("How much did you feel nervous, anxious, or on edge?" and "How much interest or pleasure did you have in doing things?"). Both items were scored on a 5-item

Likert scale (1=not at all, 2=a little, 3=somewhat, 4=quite a bit, and 5=a lot or extremely).

Weekly self-reported measures of well-being were also collected. The questions covered topics such as substance use, physical pain, connectedness to others, reception and giving of social support, general activity, and management of negative feelings. Items were scored on a 5-item Likert scale that matched the scale for the PHQ-4 and PROMIS-29 Anxiety (1=not at all, 2=a little, 3=somewhat, 4=quite a bit, and 5=a lot or extremely).

App use data were collected using the IntelliCare platform. These data contained 1 time-stamped entry per participant per app launch. Each entry included information such as the name of the app used and the launch duration in milliseconds.

4.3.5 Missingness

The rate of missing data was 39.6% among all participants (including those who dropped out at any point during the study); this rate is consistent with the often-high dropout rates in mHealth studies [16]. Among patients who completed the baseline survey, the missingness rate was 10%. Only patients who completed the baseline survey and used at least one mobile app in the IntelliCare suite were included in our final analysis (14/40, 35%).

4.3.6 Data Preprocessing, SCT Construct Grouping, and Temporal Feature Engineering

Given our overarching goal of examining the interplay between mood and engagement, we selected a theory-driven approach for grouping participants according to the *personal* construct of mood, based on a wealth of literature showing that patients with breast cancer vary with regard to their distress levels and trajectory over the course of treatment. Specifically, we grouped participants according to their baseline depression and anxiety symptoms and weekly mood [9, 11]. For symptoms of anxiety and depression, we segmented users into high and low groups according to their baseline scores. Cutoff values for determining group placement were identified using the PHQ-4 and PROMIS-29 scoring guidelines. Users who scored 3 on the PHQ-4 Anxiety subscale or who scored 60 on the PROMIS-29 Anxiety subscale were placed in the anxious group, whereas the rest were placed in the group with low anxiety. Similarly, users who scored 3 on the PHQ-4 Depression subscale or who scored 60 on the PROMIS-29 Depression subscale were placed in the group with high depression, whereas the rest were placed in the group with low depression.

We selected 2 time windows for our analysis: the entire 7-week study lifetime and 1-week intervals (e.g., week 1 and week 2). Labeling of weekly mood was conducted in a manner similar to the labeling of depression and anxiety levels at baseline. Participants with scores of 4 for weekly anxious mood were labeled anxious, and participants with scores of 2 for weekly depressed mood were labeled depressed. We note that the cutoff score for depression was applied in the inverse direction because

of the nature of the question, "How much interest or pleasure did you have in doing things?"; that is, replying 1=not at all or 2=a little indicates a depressed mood.

We conducted temporal feature engineering by hand using domain knowledge and adapting approaches from related studies. Notably, we closely followed the approach of Cheung et al. [44] to quantify the *behavioral* construct of user engagement from logs of app use data. For instance, to calculate frequency, we grouped raw app use logs by participant and period (e.g., week) and calculated the number of times the app was used during that period. We extracted 3 main measures of engagement from the raw app use data: frequency (number of launches), days of use, and duration of use. Variants of these measures (e.g., mean frequency and duration between launches) were also included in our analysis. Table 1 provides an overview of each of the 5 FSs used in the analysis.

After splitting the data into the train and test sets, we conducted multiple imputation using the MICE package [176] to handle missing values in self-reported measures. Class imbalance in the classification tasks was handled using the Synthetic Minority Oversampling Technique (SMOTE) [177], a technique that synthesizes new samples from the minority class feature space.

Table 4.2: Feature sets (FSs) used in the analysis.

FS	Description	Example features
FS1	Engagement features for all apps	Frequency of use for all apps combined, days of use, duration of use, and mean duration of use
FS2	Engagement features for only the most frequently used app or apps	Frequency of use for the app "Worry Knot" and days of use for the app "Thought Challenger"
FS3	Self-report features+engagement features for all apps	PROMIS[a] social support score, frequency of use for all apps combined, and days of use
FS4	Self-report features+engagement features for only the most-used app or apps	PROMIS social support score, duration of use for the apps "Thought Challenger" and "Worry Knot"
FS5	Self-report features+engagement features for each individual app	PROMIS physical pain score, frequency of use for the app "Worry Knot," and days of use for the app "Daily Feats"

[a]PROMIS: Patient-Reported Outcomes Measurement Information System.

4.3.7 Modeling and Prediction

Explanatory Analysis of Engagement Across Baseline Affect Groups

For each measure of depression and anxiety, we graphically analyzed the distributions of engagement measures at weekly intervals for both the low and high groups. Given the size of our data set, we analyzed engagement across all apps rather than by individual or groups of apps to avoid bias because of sparsity. Furthermore, the IntelliCare apps are conceptualized as belonging to the same intervention, and individual apps target related areas of mental health. Graphical analysis revealed notable differences in engagement between the groups with low and high anxiety and

between the groups with low and high depression.

Correlation Analysis of App Engagement and Weekly Mood

To study the correlations between app engagement metrics and weekly mood, we fit linear mixed models to account for the repeated measures within each participant, using the subject as a random effect (i.e., random intercepts) and different app engagement FSs as fixed effects. Specifically, we fit linear mixed-effects models with the least absolute shrinkage and selection operator with tuned penalty parameter and weekly anxious mood as the outcome variable on 4 FSs from Table 1 and repeated this process using weekly depressed mood as the outcome variable. Self-reported features were used as control variables.

Predictive Modeling of Weekly Mood

We wanted to investigate whether engagement with mobile apps can be used to predict weekly anxious and depressed moods, as specified in our process. We considered the case of depressed mood and formulated a binary prediction problem as follows: given a vector of a participant's app use activity and survey scores for a given week, we predicted whether the participant was depressed (1) or not depressed (0).

Binary prediction problems are well-handled by tree-based classifiers. These classifiers make decisions by splitting into one of several paths at each decision point or node. Thus, possible decision paths that can be taken to reach the final prediction are akin to the branches in a tree, with possible final predictions akin to the leaves.

Tree-based models are known for their inherent feature selection capabilities and robustness to small sample sizes, which makes them a good fit for our analysis. We selected 2 popular tree-based classifiers, XGBoost (XGB; [178]) and Random Forest (RF), and ran these with leave-one-subject-out cross-validation (LOSOCV) to predict weekly anxious mood and weekly depressed mood separately on FS3, FS5, and FS4.

The model hyperparameters were tuned using gridsearch, which attempts many combinations of different hyperparameters to find the optimal combination (i.e., the combination that produces a model with the best performance). In our case, we paired gridsearch with a variant of k-fold cross-validation called stratified group k-fold cross-validation. This technique is similar to LOSOCV in that it prevents data leakage by ensuring that no subject from the training set also appears in the testing set. It also has the additional benefit of creating stratified splits, such that the balance of positive and negative class labels (1 and 0 seconds) is roughly the same in the training set as in the testing set. This approach, similar to the SMOTE, helps mitigate the effects of class imbalance in smaller data sets.

4.3.8 Results

Explanatory Analysis of Engagement Across Baseline Affect Groups

Both the participant groups with high anxiety and high depression experienced decreases in all 3 engagement measures between week 1 and week 7, as shown in Figure

2. Notably, the groups with high anxiety and high depression started at week 1 with higher group means than their respective low group counterpoints but slowly declined across measures over time. In contrast, users with low anxiety and low depression saw gradual rises across all measures, with a sharp peak around weeks 5 to 6, followed by a subsequent decrease. Interestingly, participants with low anxiety and low depression ended the study at week 7 with approximately the same group means as their respective high group peers.

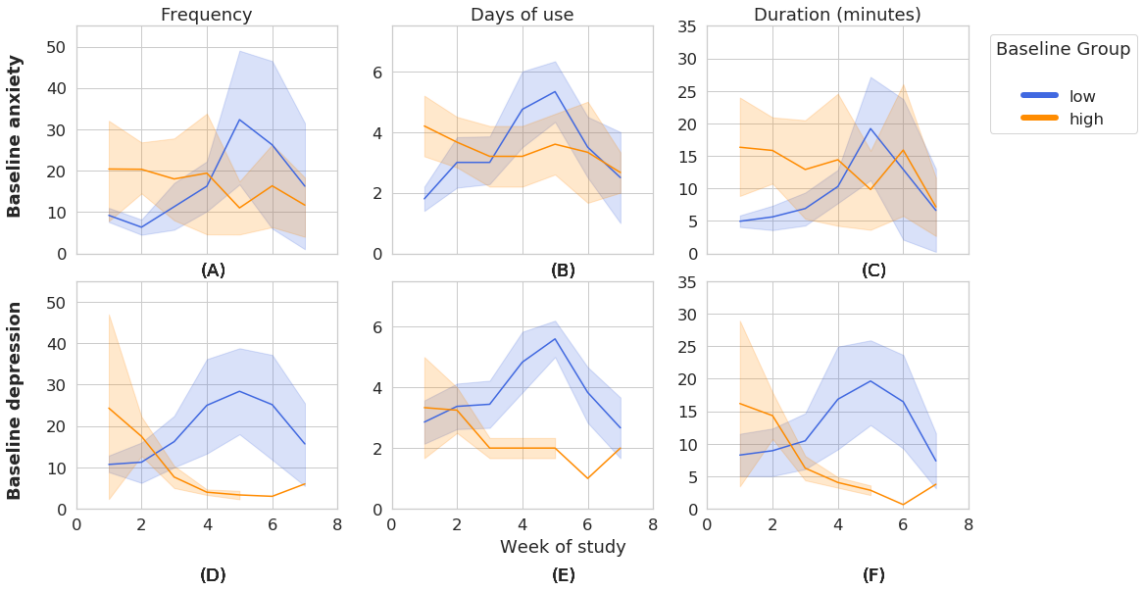


Figure 4.2: Comparison of weekly engagement metric means (with 68% CI) between 8 participants with low anxiety and 6 participants with high anxiety (A-C) and between 10 participants with low depression and 4 participants with high depression (D-F).

Correlation Analysis of App Engagement and Weekly Mood

The correlation analysis results are shown in Table 4.3. Several features of engagement provided significant correlations with weekly mood at $P<.05$. When engagement features for all apps were used (FS1), anxiety negatively correlated with the minimum duration (-0.0459). When features of only the most-used apps were used (FS2), depression negatively correlated with the week of study (-0.1826) and frequency (-0.1304) and positively correlated with days of use (0.4565), minimum duration (0.0414), and maximum duration (0.0248). The results for FSs FS3 and FS4 show that the inclusion of self-reported features as control variables improves model fit (indicated by root mean square error). When both self-report and engagement features for all apps were used (FS3), depression negatively correlated with frequency (-0.086), mean duration (-0.0637), and maximum duration (-0.0215) and positively correlated with total duration (0.0024), duration SD (0.098), and minimum duration (0.0978). Finally, when both self-report and engagement features for only the most-used apps were used (FS4), depression positively correlated with the minimum duration (0.0917) and maximum duration (0.0386). Interestingly, no significant correlations were observed between the selected app use features on weekly self-reported anxiety levels for FSs FS2, FS3, and FS4. We caution against overinterpreting this finding, given the limited sample size; rather, these results demonstrate the feasibility of identifying correlates with mood from heterogeneous data sets of engagement.

Table 4.3: Linear mixed model results stratified by feature set (FS) and outcome variable.

Outcome variable	**FS1,[a] coefficient (P value)**		**FS2,[b] coefficient (P value)**		**FS3,[c] coefficient (P value)**		**FS4,[d] coefficient (P value)**	
	Anxiety	Depression	Anxiety	Depression	Anxiety	Depression	Anxiety	Depression
Week of study	0 (—[e])	−0.16 (.14)	−0.0063 (.93)	−0.1826 (<.001)[f]	0.1122 (.22)	0.0659 (.62)	0.0643 (.43)	0.1803 (—)
Frequency	−0.0169 (.55)	−0.0632 (.14)	−0.0976 (.09)	−0.1304 (.004)[f]	−0.0438 (.12)	−0.086 (.004)[f]	−0.1747 (.001)	−0.5962 (—)
Days of use	0.0761 (.53)	−0.0737 (.74)	0.1757 (.08)	0.4565 (<.001)[f]	0.1047 (.38)	0.2374 (.25)	0.2909 (.02)	1.5607 (—)
Total duration	0.0003 (.67)	0.0021 (.12)	0.0011 (.63)	−0.0017 (.17)	0.0009 (.24)	0.0024 (.01)[f]	0.0026 (.24)	0.0009 (.68)
Mean duration	0.0237 (.17)	−0.027 (.24)	0.0071 (.78)	−0.0336 (.12)	0.0007 (.97)	−0.0637 (.03)[f]	−0.0092 (.66)	−0.1536 (—)
Duration SD	−0.0172 (.36)	0.0354 (.45)	0.0055 (.83)	−0.0093 (.66)	−0.0002 (.99)	0.098 (.02)[f]	0.0026 (.91)	0.0901 (—)
Minimum duration	−0.0459 (.02)[f]	0.032 (.37)	−0.0171 (.52)	0.0414 (.03)[f]	−0.0269 (.21)	0.0978 (.01)[f]	−0.0083 (.75)	0.0917 (<.001)[f]
Maximum duration	0.0007 (.92)	−0.0105 (.44)	−0.0047 (.70)	0.0248 (<0.001)[f]	0.0004 (.95)	−0.0215 (.05)[f]	−0.0006 (.96)	0.0386 (<.001)[f]

[a]FS1: anxiety: $\alpha = 0.1$, root mean square error 0.7396; depression: $\alpha = 0.1$, root mean square error 0.7589.

[b]FS2: anxiety: $\alpha = 0.7$, root mean square error 0.8095; depression: $\alpha = 0.1$, root mean square error 1.3954.

[c]FS3: anxiety: $\alpha = 0.1$, root mean square error 0.5128; depression: $\alpha = 0.1$, root mean square error 0.4136.

[d]FS4: anxiety: $\alpha = 0.1$, root mean square error 0.5348; depression: $\alpha = 0.1$, root mean square error 0.4547..

[e]P value was not defined.

[f]Effects with a P of <.05.

Predictive Modeling of Weekly Mood

The predictive modeling results are shown in Table 3 below. FS3, which contained survey features and overall app engagement features, achieved the highest predictive accuracy (84.6%) and yielded the best outcome measures when used with an RF classifier to predict depressed mood. FS4, which contained survey features and engagement features only from the most-used apps, achieved the second-best predictive accuracy (81.5%) when used with an XGB classifier. FS5 yielded the worst results overall, likely because of a combination of overfitting and a lack of meaningful information contained in engagement features for individual apps. Overfitting is a common issue for tree-based models applied to small data sets and occurs when the model learns the training set so well that it poorly generalizes when making predictions on the test set. We note that despite using techniques such as the SMOTE and LOSOCV, which are designed to reduce overfitting, we still struggled to mitigate this issue in our predictive task. Further investigation is warranted to determine whether a larger data set might yield better predictive results.

Table 4.4: Weekly depressed mood prediction task results.

Classifier	FS	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
RF	FS3	84.61	82.50	64.42	67.75
	FS4	83.07	73.50	72.11	72.76
	FS5	66.15	50.00	50.00	49.93
XGB	FS3	78.46	67.33	69.23	68.13
	FS4	81.53	70.81	62.50	64.54
	FS5	67.69	47.95	48.07	48.00

A feature importance graph of Shapley Additive Explanations (SHAP) scores [179]

for the top classifier and FS (i.e., RF/FS3) for depressed mood prediction is shown in Figure 4.3. Self-report features such as connectedness to others (feature *Connectedness*) and receiving support from others (feature *Receive support*) were particularly important. Engagement features such as frequency and the mean duration of use were also important. As with the results of our correlation analysis, we caution against overinterpretation of the importance of individual features, given the limited sample size.

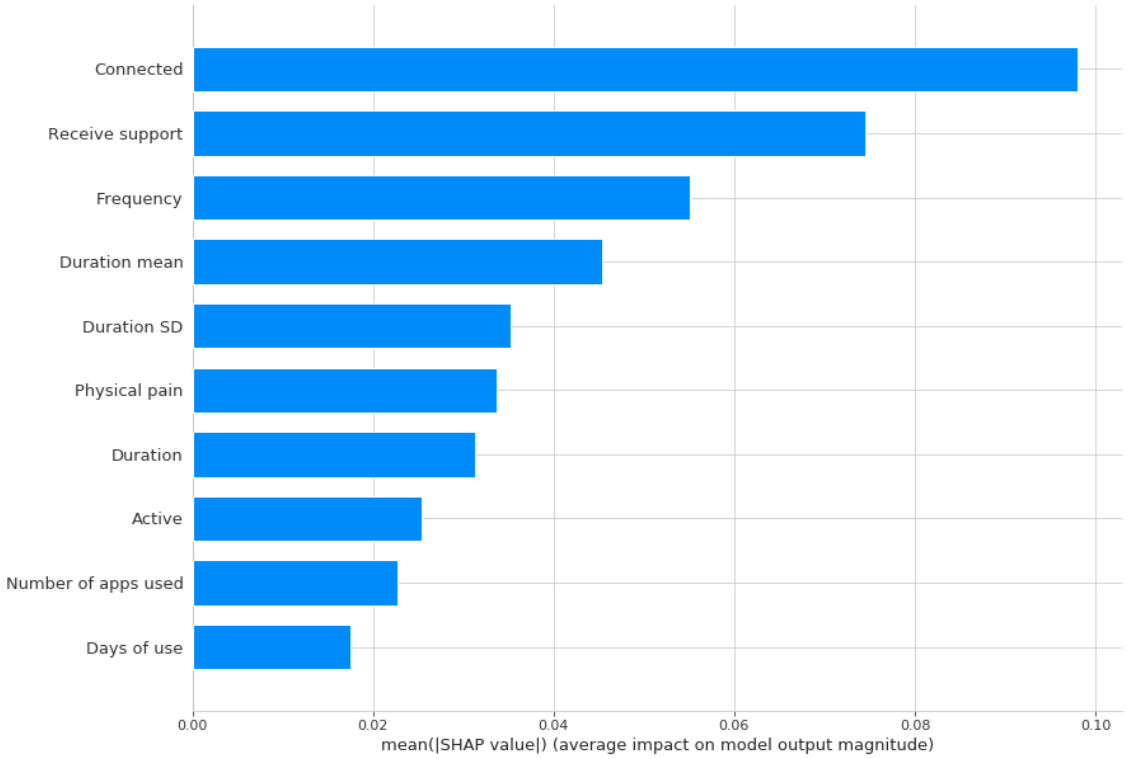


Figure 4.3: Feature importance for the prediction of depressed mood using a random forest classifier on feature set 3. SHAP: SHapley Additive exPlanations.

The findings from these exploratory analyses indicate that it may be feasible to identify the weekly moods of patients with breast cancer based on their app use

metrics.

4.3.9 Discussion

Considering the increased sophistication of mobile devices and app-delivered interventions that can capture minute details of user engagement, there is a need to develop increasingly sophisticated frameworks to make sense of user engagement data. In this study, we proposed a process for understanding the dynamic association between app engagement and mood using machine learning. Importantly, how engagement data are processed differs from study to study. The studies by Cheung et al. [44] and Pham et al. [133] drew attention to these diverse data-processing approaches and the common features that characterize engagement. Our process attempts to unify the key aspects of these approaches and refocus them on data collected from patients with breast cancer. The application of the proposed process and evaluation of statistical models support the feasibility of predicting mood status based on app engagement. The analyses and results from the case study are meant to demonstrate the potential of this approach; therefore, we caution readers not to overstate the findings of our case study. Replication of the findings in a larger data set is needed to draw more firm and generalizable conclusions.

With this caveat, the application of our process to the case study data yielded some interesting preliminary findings that may be worth pursuing in future studies. The most prominent models and theories of behavioral change highlight the importance of motivational forces to sustain a behavior [61-63], such as engagement in a mental

health app. Individuals with high levels of depression or anxiety symptoms are likely to experience low self-efficacy or a low perceived ability to perform a behavior, which is likely to result in poor engagement. Our results suggest that baseline levels of anxiety and depression affect patterns of engagement among patients with breast cancer, at least in the short term. The findings for the groups with high anxiety and high depression suggest that strong initial engagement does not necessarily lead to long-term engagement growth. In addition, the findings for the groups with low anxiety and low depression suggest that engagement may be difficult to sustain in the long term and may reach a point of diminishing returns.

The application of our process that led to the predictive results is promising in that both the RF and XGB classifiers performed well (>60% for all metrics) even with moderate amounts of data when the FS was well-curated (i.e., when FS4 and FS3 were used). This suggests that heterogeneous FSs comprising both baseline mental health measures and engagement data may be useful for predicting weekly moods when analyzed with robust classifiers. Predictions of weekly mood can, in theory, be used to personalize interventions. A dose-response relationship has been observed in digital health interventions, making it especially important to target patients when they are most open to receiving a dose of an app-delivered intervention. Heterogeneous data sets, along with high-accuracy classifiers, could be used within a just-in-time adaptive intervention (JITAI) [6] to predict the mood of patients with breast cancer. This mood could then be cross-referenced with the patient's schedule to identify the optimal time window for intervention delivery. Studies have also demonstrated that distress tends to spike in women around the time they receive an

initial diagnosis [180, 181] but that a patient's needs change throughout the course of treatment [182, 183, 184]. Such a just-in-time adaptive intervention could be further extended to learn the mood and engagement patterns of a patient with breast cancer over time and adjust the timing of the intervention accordingly. Further research is needed to determine the feasibility of implementing such interventions in vulnerable populations.

Prior studies examining the link between engagement with mHealth tools and symptoms have historically yielded mixed results; some studies have identified a direct relationship [9, 185], whereas others have identified an inverse relationship [18, 186]. Although we cannot definitively quantify this relationship in our study, both our correlation and predictive analyses suggest that paring down the available features to include only the most relevant engagement data for each individual (e.g., features from only the most-used apps) and combining self-report data with passively monitored engagement data may help researchers better identify significant predictors of mood.

4.3.10 Limitations

There are several limitations to this study that should be considered in light of these results. The results from the case study are limited in generalizability because of the small sample size. Data sparsity was a particular challenge when we attempted to break down our time windows of interest into smaller epochs, such as 4-hour windows describing different periods of the day (e.g., morning and late night); therefore, we

had to focus on daily and weekly time windows. Similar issues with sparsity occurred when we attempted to analyze the data for each individual app in the IntelliCare suite. Furthermore, our prediction task experienced overfitting. We recommend that researchers focus particularly on recruitment and retention for similar future studies to ensure that the resultant data set is sufficiently large for granular analyses.

Our study is also limited in scope as we did not account for demographic covariates, such as age, race, or socioeconomic status, in our mixed-effects model. As demographic factors are known to play an impactful role in health outcomes, we encourage researchers to include these factors in future studies on engagement with health apps. Finally, this study focused only on patients with breast cancer; therefore, our results may not be generalizable to other patient populations with cancer or other diseases.

4.3.11 Acknowledgements

This work was supported by a grant from the University of Virginia Center for Engineering in Medicine, as well as efforts supported by the University of Virginia Cancer Center. The writing of this article was supported by the National Cancer Institute (R37 CA248434; to Philip Chow) and the National Institutes of Health (T32 Training Grant in Biomedical Data Sciences; Project Number 5T32LM012416-05; to Jason Papin).

4.4 Case Study: Predicting Medication Adherence from Smart Pill Bottle Data

4.4.1 Background and Motivation

As highlighted in Chapters 2 and 3, medication adherence is a significant challenge for many patients, especially those battling diseases such as breast or lung cancer. The use of smart MEMS devices as a means of measuring adherence has grown in popularity in recent years, with smart pill management devices (e.g., Bluetooth connected pill caps) leading the charge. These devices can reliably gather the *ground truth* for an individual patient's medication-taking habits, and allow for monitoring and analysis of these habits at multiple timescales.

Consider a fictitious patient named Mary. Mary uses a smart pill bottle that records a timestamp whenever Mary opens the bottle; this timestamp indicates that Mary has taken her medication. Over the course of a week, certain habits may be made apparent based on Mary's timestamped the pill bottle data. For instance, she may take her pill around 8am and 5pm every day, which may correspond with her meal-times. On a monthly or multi-month time scale, more habits may become apparent: perhaps Mary's consistency in taking her pill is disrupted on weekends, or during certain seasons of the year. At a single time scale, Mary's smart pill cap data may not seem very useful. But for data gathered over many months and analyzed at

multiple time scales, her data becomes a rich part of her digital phenotype. This data might help Mary's doctor monitor and make adjustments to Mary's medication routine if needed. Moreover, a well-trained algorithm may be able to learn important patterns from Mary's data and even predict whether she will adhere in the next day, week, or month. This dynamic prediction of medication-taking behavior is the focus of this chapter.

A great number of studies within the mHealth literature have utilized supervised learning tasks for binary prediction of health-related outcomes. In supervised learning, data are labeled according to the true value of the outcome. Consider a scenario in which we want to predict whether a patient will be admitted to the hospital in the next 30 days from a given date. For each patient for whom we have data, we create a vector of features for that patient: clinical measures which were gathered during a previous time period (say, 15 months prior) and which were measured exactly once. I then give that vector a label: "1" if the patient was actually admitted to the hospital within 30 days of the date, or "0" if they were not admitted. The input features are then fed into an algorithm which learns from the data and predicts the label for each vector of features. This type of prediction task, in which there are no explicit temporal relationships between the features, works well in cases where features tend not to change much in value over time. But as we established in prior chapters, the human digital phenotype is a dynamic system built upon temporal relationships. Thus, traditional supervised learning approaches are insufficient for tasks such as predicting whether a person will adhere to her medication. Consequently, new methods for analyzing passively-sensed data have emerged in recent years.

Considerable attention and resources have been dedicated to measuring and predicting adherence to medications for common chronic diseases, such as heart disease, hypertension [187, 188], cardiovascular disease, and diabetes [189]. Yet, despite the enormous impact of cancer on the global population, studies of medication adherence prediction for individuals with cancer (specifically, breast cancer patients) is more rare. Moreover, existing studies of adherence prediction tend to focus on a single time scale for prediction (e.g., daily). While several works have examined predictions at multiple time scales, to my knowledge, none of these works have focused exclusively on breast cancer patients. Given the strong ties between adherence to oral chemotherapies and survival outcomes, a major opportunity exists for researchers to identify high-performing methods for predicting adherence in this population.

In this chapter, I present a case study of medication adherence prediction among breast cancer patients in the United States. This study investigates the feasibility of predicting adherence at multiple time scales from features derived using COMP-SCT.

4.4.2 Methods

Data

Data were obtained from a National Cancer Institute sponsored study (R21CA161077) of female survivors of stage 0-3 hormone-receptor-positive breast cancer (N=36). This study was approved by the San Diego State University Institutional Review Board; sharing of anonymized data with the author of this document was approved

by the Principal Investigator of the original study. In the two-month period following treatment completion, participants were asked to begin using a MEMS cap and bottle system to track their adherence to once-daily prescribed endocrine therapy. Data were collected over an 8-month period (210 weeks) and included up to three timestamps per participant, per day from the MEMS device, indicating exactly when the bottle was opened.

Problem Formulation

I formulated our problem as a supervised learning prediction task for multivariate time series data. In this scenario, a "1" indicates the person was adherent, while a "0" indicates they were not adherent.

Preprocessing

In this case study, I refer to each time the MEMS device was opened as an "event". The original dataset included approximately 8 months of daily MEMS cap readings for each participant. I first converted the data from wide-format to long format; that is, the original dataset contained one row per participant and contained all events (e.g., pill bottle opens) for that participant for the whole study period as columns. I then aligned the timestamps to their respective days within the study period. Then, I generated featuresets at two time scales: daily (featureset *study_day*) and weekly (featureset *study_week*). Finally, I used a "washout period" to exclude all data from the first month of the study. "Lead in" or "washout" periods are commonly

employed in machine-learning based prediction studies of adherence. In such an approach, all observations from a given time period (eg., the first few weeks of a study) or all observations from certain participants may be excluded from the final analysis. A lead in period guards against algorithms learning spurious patterns from early, volatile data [84]. In this scenario, the washout period accounts for the learning curve participants faced when learning to use the MEMS device, during which time they may have exhibited unstable usage patterns.

SCT Construct Grouping

Construct grouping was not applicable to this study, as the primary component of the intervention was the smart pill bottle, which had exactly one function: recording the time at which a pill was taken. Thus, all data used in this study pertains to the *behavioral* construct.

Temporal Feature Extraction

Existing data were on a daily time scale and included the exact timestamp of each event, the number of events per day, and whether a participant was adherent each day. Following Doryab et al. [5], I added several new temporal features. For the *study_day* featureset, new features included indicators of whether a given day was a weekday or weekend day and whether the MEMS device was used during a given time period on a given day (i.e., morning, afternoon, evening, or late night). For the *study_week* featureset, new features included the most common time period

during which an event occurred and the average number of events per day. For both featuresets, I also calculated commonly-used temporal metrics such as mean and standard deviation of the time of each event, as well as the mean and standard deviation of time between events.

Previous works have shown that historical adherence data is an important (if not the most important) predictor of future adherence [85, 86]. I therefore extracted lagged features of adherence and related features using a sliding window approach as in related works [85, 86, 87, 149]. Determining the optimal number of lagged features in the sliding window was important; too many lags can create a sparse dataset with low predictive power, while too few lags can obscure important temporal relationships between features. Following Gu et al.'s approach [86], I used an out-of-the-box Random Forest classifier to determine an appropriate number of lags. For each feature set, I varied the number of lags (range $[1, 7]$) as well as the *max_depth* parameter (range $[0, 5]$) as inputs into the classifier, then selected the optimal number of lags by visual inspection of classifier performance. I examined both lags and *max_depth* in tandem in order to to determine an appropriate upper bound on the *max_depth* parameter during tuning. Graphical results for *max_depth* values of 1 and 2 are shown as examples in Figure 4.4.

As Gu et al. [85] and Gu et al. [86] note, adherence prediction researchers are most interested in participants who do *not* adhere, as these are the patients most in need of clinical intervention. Therefore, specificity is of greater interest in adherence prediction studies than more commonly-used metrics such as accuracy or sensitivity. I based my final decision for the optimal number of lags on the balance between

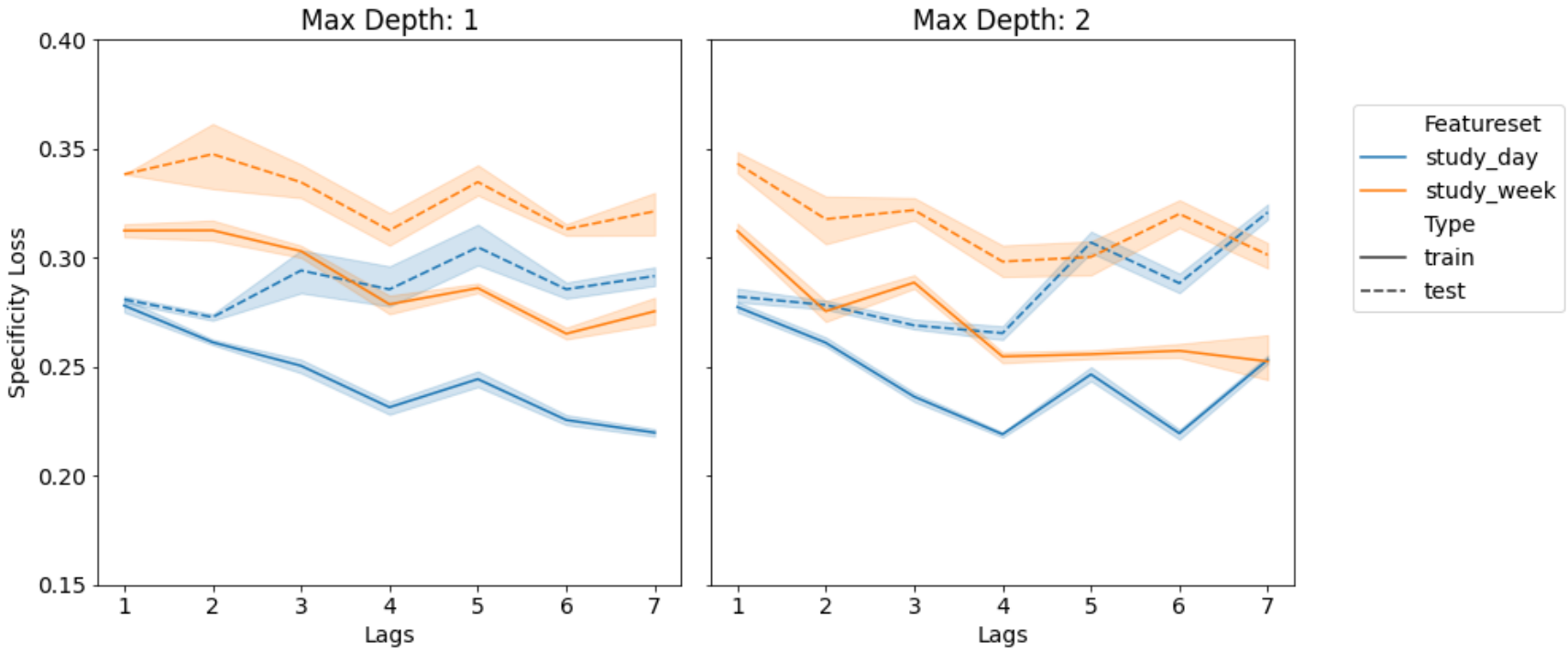


Figure 4.4: Specificity loss for lag tuning exercise, with *max_depth* of 1 (left) and 2 (right).

overfitting and *specificity* loss. This exercise led me to select $n = 4$ lags for both next day and next week prediction [1]. Given that performance was best using very shallow trees, I also set an upper bound of 3 for the *max_depth* parameter in the search grid prior to beginning the prediction task.

Prediction tasks

I formulated two unique prediction tasks: one for next day adherence prediction (Task 1, using featureset *study_day*) and one for next week adherence prediction (Task 2, using featureset *study_week*), respectively. Daily adherence was defined as taking the medication (i.e., opening the MEMS device) within 24 ± 6 hours of taking the last dose Weekly adherence was defined as being adherent for at least 80% of the time. [2]

I conducted 5-repeated, nested 5-fold cross validation with hyperparameter tuning. I split data at the subject level, which guards against data leakage by ensuring that subjects who appear in the training set do not appear in the test set. After splitting the data, I imputed missing numeric features using sklearn's IterativeImputer for multiple imputation and imputed nominal features using mode imputation. I also one-hot-encoded nominal features. For classifiers which are sensitive to scaling (e.g., SVM and Logistic Regression), I scaled features to be in the range $[0, 1]$ using sklearn's MinMaxScaler. Following Gu et al. [85] and Gu et al. [86], I used a Random Forest classifier for feature selection. I tuned hyperparameters using the ray-tune python package's implementation of gridsearch []. Importantly, I used sensi-

[1] It should be noted that selecting the optimal number of lags based on visual inspection is naturally subjective. For instance, n=2 lags might have been a more optimal choice for next-day prediction (Task 1). However, this choice would have made it difficult to perform a fair comparison between Task 1 and Task 2. I selected 4 lags for both tasks to ensure a fair comparison between the tasks for each classifier, and to ensure that the data had a large enough lookback window for each observation to generate meaningful predictive insights.

[2] An adherence threshold of 80% is standard within the medication adherence literature, though whether it should remain the standard is up for debate [190].

tivity as the gridsearch scoring metric so as to select the best classifier for identifying patients who would *not* adhere.

I compared the performance of fmywell-known machine learning classifiers on each feature set: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and XGBoost (XGB; [178]). I calculated the mean and standard deviation across all runs and folds for each evaluation metric. Results for Task 1 (next day prediction) and Task 2 (next week prediction) are shown in Table 4.5.

4.4.3 Results

Task 1 was associated with better model performance and stability across all models, highlighting the importance of capturing frequent, granular measures of adherence during deployment studies. XGBoost (XGB) achieved the highest mean accuracy (87.26%) and sensitivity (93.97%), while the Random Forest (RF) classifier achieved the highest specificity (74.03%) and mean AUC (0.86). For Task 2, the Random Forest classifier achieved the highest mean accuracy (75.13%), sensitivity (81.01%) and AUC (0.81) Interestingly, the Logistic Regression (LR) model achieved the highest specificity in Task 2 (71.18%) and achieved relatively high specificity in both tasks in comparison to the other models. This finding warrants further investigation, given that tree-based models tend to perform best with tabular data especially in medical contexts. However, such an investigation is outside the scope of this current study.

It is worth noting that, like Gu et al [85], I observed the unusual occurrence of higher performance on the test set than on the train set for select models and metrics,

Table 4.5: Results for adherence prediction for Task 1 and Task 2

Task 1: Next day Adherence Prediction						
Method	**Accuracy (%)**	**Precision (%)**	**Sensitivity (%)**	**Specificity (%)**	**F1 Score**	**AUC**
LR	83.98 ± 0.40	90.51 ± 0.13	87.61 ± 0.66	73.49 ± 0.54	0.89 ± 0.00	0.85 ± 0.05
RF	84.38 ± 0.46	90.72 ± 0.11	87.97 ± 0.79	74.03 ± 0.55	0.89 ± 0.00	0.86 ± 0.04
SVM	86.38 ± 0.76	88.97 ± 0.19	93.21 ± 0.93	66.65 ± 0.35	0.91 ± 0.01	0.80 ± 0.06
XGB	87.26 ± 0.09	89.42 ± 0.14	93.97 ± 0.30	67.90 ± 0.56	0.92 ± 0.00	0.85 ± 0.05
Task 2: Next week Adherence Prediction						
Method	**Accuracy (%)**	**Precision (%)**	**Sensitivity (%)**	**Specificity (%)**	**F1 Score**	**AUC**
LR	74.36 ± 0.47	77.13 ± 0.71	76.88 ± 0.87	71.18 ± 1.36	0.77 ± 0.00	0.80 ± 0.05
RF	75.13 ± 0.84	76.02 ± 0.65	81.01 ± 1.23	67.69 ± 1.03	0.78 ± 0.01	0.81 ± 0.05
SVM	75.06 ± 0.85	76.47 ± 1.47	80.00 ± 0.88	68.82 ± 2.86	0.78 ± 0.00	0.79 ± 0.06
XGB	74.59 ± 1.30	75.83 ± 1.02	79.96 ± 1.39	67.79 ± 1.29	0.78 ± 0.01	0.80 ± 0.05

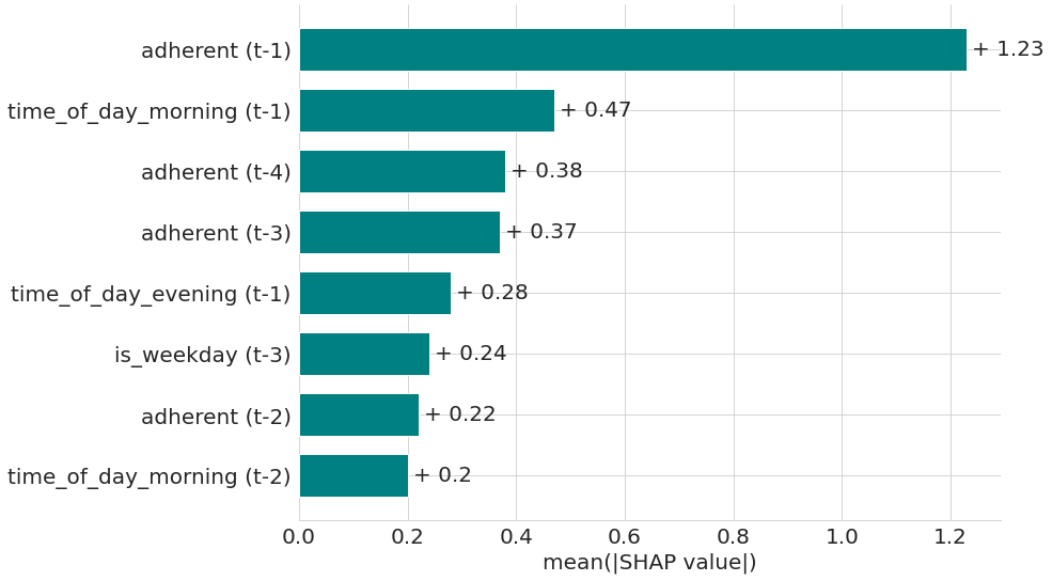
indicating that the model behavior is not fully explainable. This situation points to a need for better tools for explainable AI, an important topic that is beyond the scope of this paper.

As in Case Study 1 (Section 4.3), I calculated the feature importance using Shapley Additive Explanations (SHAP) [179]. Figure 4.5 shows the magnitude of the impact of the most important features on the model performance, for the top model for each task. I calculated the importance of each feature as the mean SHAP value across all runs and folds. Consistent with prior literature on adherence prediction, recent prior adherence played an important role in predicting future adherence. For Task 1 (next day prediction; Figure 4.5a), adherence for the previous day was the most important feature. The time of day during which the patient took their medication the previous day was also important. Interestingly, the distinction between weekday days and weekend days was important as well, but only for the day three days prior. This finding raises the possibility that participants were more likely to be adherent on a Monday, the first day of the week.

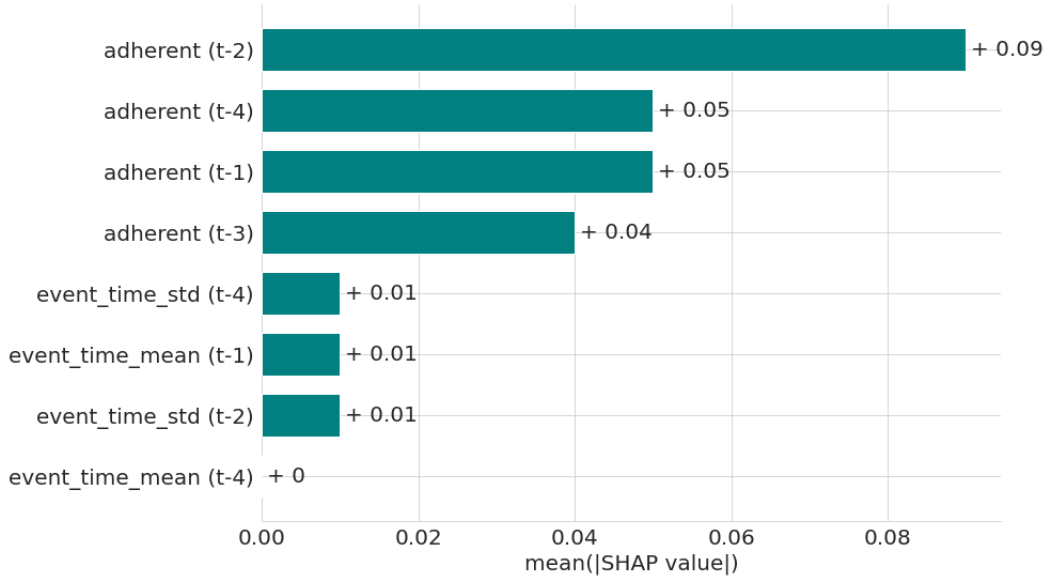
Features for Task 2 (next week prediction; Figure 4.5a) had a much lesser impact on Task 2 model performance than did the features for Task 1 on Task 1 model performance. For Task 2, adherence in the second most recent week was most important. Features such as the average time at which the patient took their medication each day as well as the standard deviation for that time were also slightly important, though their impact was small in comparison to that of other features.

4.4.4 Discussion

The current study investigates the feasibility of predicting future medication adherence from passively-sensed data from smart MEMS devices. The best model's performance is comparable to the performance of leading adherence prediction mod-



(a)



(b)

Figure 4.5: Feature importance for the top classifier for each prediction task. Figure 4.5a shows the feature importance for XGBoost, the top-performing classifier for Task 1 (next day prediction). Figure 4.5b shows the feature importance for Random Forest, the top-performing classifier for Task 2 (next week prediction).

els [84, 85, 86], and the model even outperforms similar models for multiple evaluation metrics (e.g., accuracy, sensitivity, and AUC) despite the smaller sample size. Model performance is dependent on a number of factors including the nature of the problem domain, the prevalence of the outcome variable, and the population sample size [191], therefore, I cannot claim that this study's best model is inherently "better" than existing models. Rather, our results demonstrate that predicting adherence at multiple timescales is feasible, and shows the usefulness of COMP-SCT in guiding the feature engineering and model optimization processes in supervised learning tasks for clinical outcomes prediction. Specifically, my addition of features related to the time of day in Task 1 and the time of the week (i.e., whether a day was a weekday) for Task 2 played an important role in the model's performance, as shown in Figure 4.5.

Prior works have shown that the relationship between features of human behavior and clinically-relevant outcomes such as mood can vary by the granularity of the time scale on which the input features were measured (e.g., time of day, time of year, etc.) [5]. This phenomenon complicates the nature of adherence prediction in real-life settings. Clinicians, who are already very busy, are highly unlikely to be able to respond to daily updates about their patients' adherence patterns. Moreover, as illustrated by the example of the fictitious patient Mary at the beginning of this chapter, limiting adherence prediction to a single time scale risks missing patterns inherent at smaller time scales. In practice, multi-timescale approaches to adherence prediction will need to adapt to existing clinical practice, including triage. In the context of medication adherence monitoring, "triage" in this case entails using automated messages to provide support to the patient when they are at low risk,

and informing nurses on duty or the patient's doctor if they are at higher risk. For instance, a patient might be considered at low risk if they have not yet missed a dose of a certain medication, but they are predicted to be at risk of missing a dose the following day. In this case, Toscos et al.'s approach [73] could be adapted to securely deliver a support message to the patient via a mobile app or EHR portal, but *before* the missed dose occurs.

Importantly, this study's most accurate model has a high sensitivity and thus can properly capture most cases of nonadherence on a daily time scale. This particular result is promising, given the challenging nature of risk detection and adverse outcome prevention in clinical settings. If I constructed a final XGBoost model based on my approach in this study and deployed it in a clinical environment, it could be reliably be used to detect nonadherence among breast cancer patients and help guide decisions about when and how frequently to reach out to breast cancer patients at risk of not adhering. The results from this study thus help lay a foundation for adherence monitoring and prediction in certain oncology contexts. Further, these results contribute to domain knowledge of key behavioral factors that influence medication adherence among breast cancer patients, a population whose adherence patterns are not yet fully understood.

4.4.5 Limitations

Several limitations exist in this study. Unlike other studies on medication adherence, this study does not examine factors such as the total quantity or days supply for each

dispensed medication, which may be important factors [190]. This study also does not leverage additional data sources such as EHR or medical claims data, which have been shown to achieve fair predictive performance for adherence patterns [192].

I impose a strict 80% threshold for adherence; however, related work by Lo-Ciganic et al. has shown that an 80% threshold is not always predictive of adherence and that this threshold may vary by disease and medication [190]. Additionally, classifier performance is often dataset-specific [191]; therefore, this study's classification approach might not yield equally-good results if applied to another dataset of medication adherence. Further research is needed to determine whether a different threshold for adherence should be used, and whether this study's approach can be generalized to achieve good predictive results on similar datasets.

Chapter 5

Conclusions and Future Work

5.1 Summary of Contributions

In **Chapter 1**, I presented a review of the relevant literature on user context, which included an examination of leading scholars' definitions of "context", the role context plays in characterizing an individual's health status via digital phenotyping, and the open challenges to user context detection *in situ*. I also reviewed the prominent frameworks that laid the foundation for modern digital phenotyping studies, examined the complex relationship between engagement with digital phenotyping interventions and subsequent attrition, and discussed current state-of-the-art approaches to predicting medication adherence from passively-sensed data using machine learning.

In **Chapter 2**, I discussed our work on the Multiscale Monitoring and Intervention (MMI) system, which is designed to guide the design, deployment, and improvement of mHealth interventions for medication adherence by utilizing passively-sensed indicators of user context. This chapter argued that adherence is a challenge best characterized by constructs from Social Cognitive Theory (SCT), and provided an

explicit mapping from SCT constructs to sensors and contextual features of adherence. Our presentation of the MMI system represents a twofold contribution: We laid the foundation for grounding future interventions for medication adherence in SCT, and we set the stage for developing and deploying such interventions in the wild.

In **Chapter 3**, I presented our next major step in making the MMI system a reality: A qualitative study of lung cancer patients' perceptions of and needs for symptom and adherence monitoring interventions. This study revealed key insights about patients' preferences, needs, and priorities, such as the need to receive unobtrusive notifications on one's own schedule and the desire for adequate support resources from the study team during the intervention period. Importantly, this study also presented several actionable implications for design, such as increasing patients' agency over their data and notifications.

In **Chapter 4**, I presented COMP-SCT, a framework for deriving personal, behavioral, and environmental features of user context at multiple time scales. COMP-SCT make several novel contributions to the field of mHealth. It maps SCT constructs to low-level physiological data and higher level features of medication adherence and facilitates the analysis of SCT constructs at multiple time scales. Further, it helps recontextualize medication adherence as a system, which takes into account not only the user (the patient), but also the environmental factors that impact the user (e.g., how much interaction the user has with friends and clinicians). This approach is novel considering that similar frameworks tend to take a more user-centric view of challenges such as adherence, thus missing a critical piece of the user's context: their

environment. I also described our application of COMP-SCT to two case studies of breast cancer patients, demonstrating its utility for deriving *personal* and *behavioral* features of user context and for driving actionable clinical predictions at the level of the current state-of-the-art.

5.2 Future Work

The work presented in this dissertation lays the groundwork for future studies of medication adherence and the factors that contribute to adherence among oncology populations. Several viable directions for future work exist, which I discuss briefly in the following sections.

5.2.1 Mobile Health (mHealth) Devices and Remote Monitoring: Ongoing Work with Lung Cancer Patients

In Chapter 3, we presented our early work in understanding patients' attitudes toward smart devices for managing oral chemotherapy during lung cancer treatment. The natural extension of this work is to pilot mHealth technologies with real patients. Importantly, mHealth devices may serve as a complement to traditional clinical trials for oncology, given that devices can easily capture patient reported outcomes (PROs) such as sleep quality, medication side effects, and mental wellbeing in real time and may lead to better overall outcomes for patients whose physicians closely monitor their reported progress [80]. The advantages of mHealth devices in oncology

care have led to a number of clinical trials on the use of such tools for medication and side effect management. Mauro et al conducted a randomized controlled trial to evaluate the effect of a smart pill bottle among patients with multiple myeloma who were taking oral chemotherapy [112]. The study followed 40 patients for 6 months, with 32 in total completing the study. The intervention group received smart bottles with visual and audio reminders activated (e.g., blinking lights, audible "chimes") as well as support and monitoring from a pharmacist. The control group received the same smart bottles as the intervention group, but these bottles did not have reminders turned on. Results showed that adherence was significantly higher in the intervention group, suggesting that smart technologies such as pill bottles may help improve adherence when reminders are turned on and clinicians are able to follow up with patients if needed. Further, Greer et al. conducted an 18-month randomized controlled trial of a mobile app with 181 cancer patients prescribed oral anticancer medications [142]. The app prompted participants to take medications and track symptom severity and included Fitbit integration. They found that participants in the intervention group who struggled with anxiety and adherence at baseline showed higher MEMS-measured adherence rates over the study duration.

As discussed in Chapter 1, several preliminary studies on remote monitoring for lung cancer care using smart devices have emerged in recent years. However, comparatively few clinical trials focused exclusively on symptom and adherence monitoring for lung cancer have taken place. Moreover, few trials have focused exclusively on *oral tyrosine kinase inhibitors (TKI)*, a type of chemotherapy commonly used in patients with non-small cell lung cancer (NSCLC) with EGFR mutations or ALK gene rear-

rangements. Perhaps the most relevant non-clinical study in the area of mHealth for lung cancer and TKIs is that of Pereira-Salgado et al. [114]. The authors conducted a 10-week pilot test of an mHealth system for medication- and symptom-management with 9 patients with chronic myeloid leukemia (CML). The system was designed to help patients better adhere to their oral tyrosine kinase inhibitor (TKI) medications by sending medication reminders via text, asking patients to record their side effects via an online web portal, and providing self-care tips. Patients reported feeling more informed about their care (including care of side effects), and both patients and clinicians and found the system helpful for establishing healthy medication-taking routine. These results are a promising step toward mHealth interventions for lung cancer, but more work must be done to fully understand the role of mHealth devices in symptom and adherence monitoring.

Tie in the contextual (mobile) data more clearly. What do we expect to learn from usage patterns (e.g., time of day)? From self-reported symptoms?

We have begun the next phase of our project on symptom and adherence monitoring among lung cancer patients, in collaboration with oncologists at the NCI-designated Emily Couric Cancer Center at the University of Virginia. In this phase we leverage Sensus, a novel mobile sensing system for Android and iOS capable of collecting data from smartphone and Bluetooth-enabled sensors and EMAs on a schedule, randomly, or using sensor-triggers [32]. We also leverage Swear, a mobile app for smartwatches developed by the project team to collect multimodal signals related to health [193]. Using the Sensus and Swear apps on smartphones and smartwatches, respectively, we capture a rich dataset containing key elements of user context for each of the

Social Cognitive Theory constructs. We ask patients to self-report personal factors such as the frequency and intensity of their symptoms and the quality of their sleep, behavioral factors such as whether they took their medication, and environmental factors such as whether they are with another person. We also passively capture objective measures of user behavior in relation to their environment. For instance, we capture movement patterns via accelerometer and gyroscope, which can indicate activities like exercise. After the study's completion, we will extract relevant features using COMP-SCT and will leverage deep learning models to model side effect and adherence patterns.

The results of our study will help determine the feasibility of conducting a larger-scale, randomized-control trial. This trial would evaluate whether the MMI system can lead to meaningful changes in patient outcomes, including more timely AE detection, reduced AE symptom burden, and reduced treatment interruption and/or discontinuation. The results of this study will also be used to develop a clinical dashboard for remote, real time symptom monitoring. This dashboard will enable clinicians to intervene earlier when a patient is experiencing side effects, thus reducing severity of symptoms, improving quality of life, and avoiding drug discontinuations and dose reductions.

5.2.2 Predicting User Behavior Using Both Static and Temporal Data

The case studies I presented in Chapter 4 demonstrate the feasibility of predicting personal and behavioral outcomes from dynamic (temporal) data. However, the nature of mHealth studies is such that dynamic data rarely exists in isolation. Often, researchers will administer baseline assessments prior to the start of an intervention in order to gather demographic information and information about an individual's level of health at a single point in time (e.g., at the start of a study). Similarly, electronic health records will contain a single date per diagnosis, given that it makes no sense to be diagnosed with the same disease more than once. Baseline assessments and select portions of EHR data collectively represent the rich *static* data that occurs alongside dynamic, passively-sensed data in real world contexts. Recent work has shown that complex models such as neural networks can model and predict static health data with very high predictive accuracy [194]. Moreover, multiple studies have shown that these complex models can be used to model static and dynamic data together, also with high predictive accuracy [195, 196, 197]. For instance, Esteban et al. developed a novel modeling approach in which static data is processed with a feedforward neural network and dynamic data is processed with an recurrent neural network (RNN). The authors applied their approach to a dataset of 2061 patients with kidney failure, using several variants of RNNs (such as Long Short-Term Memory models, or *LSTMs*, and gated recurrent units, or GRUs). The authors compared these variances to Random Forest and Logistic Regression, and ultimately showed that the GRU yielded the

best AUC (0.833 ± 0.006) when predicting all endpoints (i.e., transplant rejection, transplant loss, and death) across all time horizons (6 months and 12 months).

As Leontjeva and Kuzovkin point out, static and dynamic data are rarely used together because traditional (i.e., supervised learning) classifiers are not able to handle such combinations of data [196]. Moreover, as evidenced by the literature, few studies have combined static and dynamic data in health-related contexts. A novel extension of the work we presented in Chapter 4 would be to utilize baseline demographic data and measures of health (e.g., scores from assessments of depression such as the PHQ-9) alongside the passively sensed sensor data, either through the combination of a feedforward neural network and RNN as in [195], or through the use of a hybrid model as in [196]. Such an approach would yield deeper insights into how multiple SCT constructs influence patients' adherence decisions. Consequently, researchers and clinicians would be able to more efficiently determine how to personalize adherence interventions based on an individual's background or comorbid conditions.

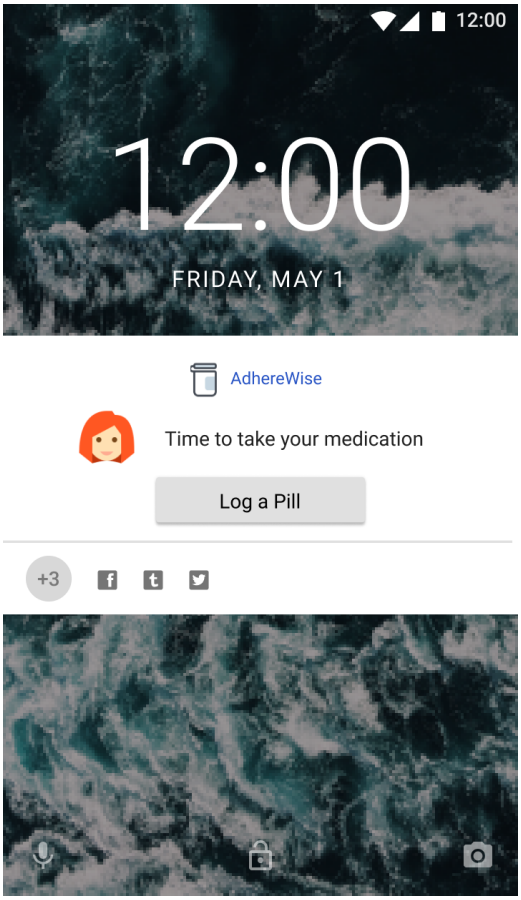
5.2.3 Design of Just-In-Time Adaptive Interventions for Adherence and Symptom Monitoring

The interview study presented in Chapter 3 as well as the case studies presented in Chapter 4 have advanced our knowledge of factors that contribute to human behavior in oncology populations. But these studies are only the first step toward the end goal of mHealth research: improving patient health outcomes. The findings in this dissertation could be used under the guidance of the MMI and COMP-SCT

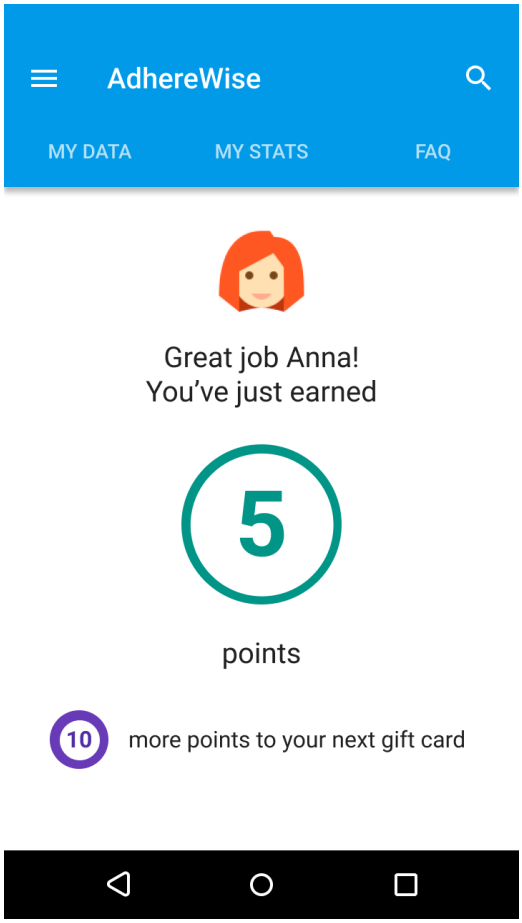
frameworks to inform the design of multi-scale *Just-in-Time Adaptive Interventions (JITAIs)* for medication adherence. JITAIs were first formally proposed by Spruijt-Metz and Nilsen in 2014 as a mechanism for leveraging health behavior theory and context representation to deliver personalized interventions at the moment they are needed (Spruijt-Metz and Nilsen 2014). JITAIs are designed to both sample user context in the moment and deliver an appropriate intervention based on that context. They are valued for their ability to "capitalize on periods of heightened susceptibility to positive health behavior changes". JITAIs are predicated on several core principles: namely, that a rich representation of user behavior and the context of that behavior via mobile devices is feasible; and, that representation of context, both past and present, is "dynamic knowledge" that can be used to inform real-time behavior change interventions. While JITAIs do not solve all the problems associated with capturing of user context, they help to break down both context and user goals into smaller pieces. Spruijt-Metz and Nilsen provide several examples of how a this might work in practice; for instance, GPS data could be used to nudge an individual towards a healthy eatery as they walk to get lunch, helping to encourage long term behavior change toward healthy eating. As with all interventions that target behavior change, JITAIs can only be successful if the user is both open and receptive to receiving in-the-moment support. Recontextualized in another way, JITAIs initially target the "Contemplation" and "Determination" phases described in the well-known Transtheoretical (stages of change) Model [198]. A successful JITAI will then gradually adapt to support the action and maintenance phases as the user develops healthier habits over time.

JITAIs have been proposed and applied to a variety of health challenges, including addiction [199], smoking cessation [200], stress [147, 201, 202], and physical activity [203]. Few JITAIs exist, however, for the promotion of medication adherence (with a study of WisePill as a notable exception [204]). The work I have presented in Chapter 4 (Case Study 2) could readily be used to implement the behind-the-scenes predictions that would drive a medication adherence-focused JITAIs forward. As an example, let us consider the case of a simple JITAI for adherence prediction in breast cancer patients. When designing a JITAI, one must consider both *proximal* (short-term) and *distal* (long-term) outcomes [6]. Therefore, let us define an example distal outcome O_d as *80% adherence 3 months into the intervention period*. Now, let us define an appropriate related proximal outcome O_p: as *minimizing intervention fatigue*; note that this proximal outcome, which has been adapted from Nahum-Shani et al.'s example for a JITAI for weight loss [6], prioritizes engagement by reducing user burden. Having defined our proximal and distal outcomes, we must now consider the four key components of a JITAI: *decision points*, *tailoring variables*, *intervention options*, and *decision rules*. We might also choose to incorporate elements of gamification which the literature has identified as helpful, such as avatars and point-based incentives; let us include these as part of our decision rules. Table 5.1 below shows three example decision rule workflows for our proposed JITAI (similar to Table 2 of [6]). These include elements of *gamification*, in which components such as avatars, points, and small monetary incentives are used to encourage sustained participation.

The first and second rules makes use of an avatar which delivers an encouraging



(a) Example of avatar-delivered EMA prompt



(b) Example of points-based incentive system

Figure 5.1: Example of user interface for a JITAI to improve medication adherence. Images courtesy of Material Design Kit and Icons8.

Table 5.1: Example of a decision rule workflow set for a gamified JITAI for medication adherence

Decision Rule	Decision Point	Tailoring Vars	Intervention(s)
At 12pm If number of medication-taking events == 0: IO = [EMA prompt from clinical care team avatar to encourage recording of medication-taking event] Else IO = [provide nothing]	Specific Time of Day	Number of medication-taking events (measured by EMA accelerometer / MEMS cap)	Ask about medication-taking event OR do nothing
Let T_w be the user's average wake up time and T_s be the user's average bedtime, on a 24 hour time scale Every 6 hours If time $> T_w$ and time $< T_s$ IO = [EMA prompt from clinical care team avatar to encourage recording of medication-taking event] Else IO = [provide nothing]	Pre-specified time interval	Wake time and sleep time	Ask about medication-taking event OR do nothing
When user records medication-taking via EMA IO = [Award 1 point and show progress toward badge] Else IO = [provide nothing]	Time of EMA completion	Most recently-completed EMA	Award points OR do nothing

message to record one's medication. This approach could also be extended to use conversational agents, as Schroeder et al. did with the Pocket Skills app for Dialectic Behavioral Therapy [205]. Finally, the third rule allows for the accumulation of points for completing EMA questions. Points could then be put towards a monetary incentive such as a gift card. Figure 5.1 shows an example of a user interface that could accompany these decision rules, with Figure 5.1a corresponding to the first two rules and Figure 5.1b corresponding to the third.

5.3 Concluding Remarks

In this dissertation, I have presented a systems-level approach to mobile health focused on modeling user behavior in context. This work draws upon health behavior theory and the ever-growing body of literature on user context detection to demonstrate how personal, behavioral, and environmental factors shape our health and wellbeing. The results of our mixed-methods work with lung cancer patients and our case studies that utilize the COMP-SCT framework have exciting implications for the future of symptom and adherence monitoring in daily life, and advance the state-of-the-art for translating raw physiological data into actionable clinical insights. Though the work contained in this document is a relatively small contribution in the field of mHealth, it is nevertheless a meaningful contribution toward building user-informed, theory-grounded interventions for medication adherence and related challenges in personal health management.

Appendices

Appendix A

User Interview Questions

A.1 Smartphone App

1. In general, how did you feel about the experience of filling out the surveys on the smartphone?

2. What did you like most/least about using this smartphone app?

3. How easy or difficult was it to answer the questions in the app? (PROBE: What made it easy? What made it difficult?)

4. How would you feel if we were to ask you to use this smartphone app for a long period of time, like 6 months? (PROBE: Why would you feel like that?)

5. Would you find it easy or difficult to use the smartphone app on your own phone? (PROBE: Why would you find it easy/hard?)

6. How often would you be able to fill out surveys like this on your smartphone? (PROBE: How many days (i.e., every day, every couple of days; PROBE: How many times per day (i.e., 1 time each day, 3 times each day)?)

7. About how many questions like this would you be able to fill out at a time?

8. Are there times of the day / days of the week where it would be easier / harder / impossible for you to answer survey questions like these? (PROBE: Why would it be easier / harder / impossible?)

9. How do you feel about the notifications from the app? (PROBE: What kind of notifications you prefer: beep, buzz, or another sound or vibration?; PROBE: Would you feel comfortable having the notifications show up on your phone, in public? Why/why not?)

10. Do you have other thoughts that you would like to share about the smartphone app?

11. Do you have any other thoughts that you would like to share about filling out surveys like this on your smartphone?

A.2 Smartwatch App

A.2.1 For Participants Who Own a Smartwatch

1. How would you feel if we were to ask you to wear a smartwatch for our project every day for a long period of time, like 6 months? This would be a different smartwatch from the one that you already have. (PROBE: Why would you feel like that?)

2. Would you find it easy or difficult to remember to wear the project smartwatch every day? (PROBE: Why would you find it easy / difficult?)

3. Would you find it easy or difficult to remember to charge the project smartwatch at night? (PROBE: Why would you find it easy / difficult?)

4. Tell me about any times when you would not want to wear the project smartwatch? (PROBE: Why wouldn't you want to wear it then?)

A.2.2 For All Participants

1. In general, how did you feel about the experience of filling out the surveys on the smartwatch?

2. What did you like most/least about using this smartwatch app?

3. How easy or difficult was it to answer the questions in the app? (PROBE: What made it easy? What made it difficult?)

4. How would you feel if we were to ask you to use this smartwatch app for a long period of time, like 6 months? (PROBE: Why would you feel like that?)

5. Would you find it easy or difficult to use the smartwatch app on your own phone? (PROBE: Why would you find it easy/hard?)

6. How often would you be able to fill out surveys like this on your smartwatch? (PROBE: How many days (i.e., every day, every couple of days; PROBE: How many times per day (i.e., 1 time each day, 3 times each day)?)

7. About how many questions like this would you be able to fill out at a time?

8. Are there times of the day / days of the week where it would be easier / harder / impossible for you to answer survey questions like these? (PROBE: Why would it be easier / harder / impossible?)

9. How do you feel about the notifications from the app? (PROBE: What kind of notifications you prefer: beep, buzz, or another sound or vibration?; PROBE: Would you feel comfortable having the notifications show up on your smartwatch, in public? Why/why not?)

10. Do you have other thoughts that you would like to share about the smartwatch app?

11. Do you have any other thoughts that you would like to share about filling out surveys like this on your smartwatch?

A.3 Smart Pill Bottle / Box

A.3.1 Medication-Related Questions

1. How many prescription medicines do you currently take each day?

2. How many vitamins or supplements do you currently take each day?

3. What times of the day do you take your medicines and supplements?

4. Where do you keep your medicines and supplements (e.g., purse, night stand, medicine cabinet)?

5. In what type of container do you keep your medicines and supplements? (e.g., in a pill keeper, in the original bottle)?

6. Do you have any medicines that come in a blister pack or some other type of dispenser?

7. Do you have any medicines that are liquid?

8. Do you have any medicines that need to be refrigerated? (If yes: Are those pills or liquid?)

9. Do you sometimes have a different container for your medicines and supplements, like a smaller container that you take on a trip or an outing?

10. Does anyone help you with managing or taking your medicines? (PROBE: Who helps you?; PROBE: How do they help you?)

A.3.2 Questions About Prior Technology Use

1. Have you ever seen a smart pill bottle / box like this? (PROBE: If so, was it a box or bottle?)

2. Have you ever used a smart pill bottle / box like this? (PROBE: If so, do you still use it now? PROBE: If so, was it a box or bottle? PROBE: Do you know what brand it was/is?)

3. What did/do you think about the smart pill bottle / box that you used? (PROBE: Did you like or dislike it? Why?)

4. Do you have a smartphone? (PROBE: What type?)

5. Have you ever used an app on your phone to manage your medicines? (PROBE: If so, which app(s)?; PROBE: Do you still use it?; PROBE: How often do you use it?)

6. What do you/did you think about the app that you use to track your medicines? (PROBE: What features of the app are/were helpful?; What features of the app are/were not helpful?

7. Why did you stop using it?

A.3.3 Questions about the Pill Bottle / Box

1. How would you feel if we were to ask you to use this smart pill bottle / box for a long period of time, like 6 months? (PROBE: Why would you feel like that?)

2. Would it be easy or difficult to fit using this smart pill bottle / box into your everyday life? Why?

3. Would you be able to still use this pill bottle / box if you take multiple medicines each day?

4. How would you handle using this smart pill bottle / box if you had to be away from the place where you usually take your medicines, like if you had to take a trip away from home?

5. How do you think using this smart pill bottle / box would affect the routine that you have for taking your medicines?

6. Do you think that using this smart pill bottle / box for a long time would make it more or less likely that you would forget your medicines?

7. Do you think that it would be easy or difficult to keep track of this smart pill bottle / box? (PROBE: What would make it easy? PROBE: What would make it difficult?)

8. Do you have other thoughts that you would like to share about the smart pill bottle / box?

Bibliography

[1] The Pew Research Center, "Mobile fact sheet," 2021.

[2] G. Coppersmith, M. Dredze, and C. Harman, "Quantifying Mental Health Signals in Twitter," in *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, (Baltimore, Maryland, USA), pp. 51–60, Association for Computational Linguistics, 2014.

[3] G. Coppersmith, M. Dredze, C. Harman, and K. Hollingshead, "From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses," in *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, (Denver, Colorado), pp. 1–10, Association for Computational Linguistics, 2015.

[4] M. Boukhechba, J. Gong, K. Kowsari, M. K. Ameko, K. Fua, P. I. Chow, Y. Huang, B. A. Teachman, and L. E. Barnes, "Physiological changes over the course of cognitive bias modification for social anxiety," in *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, (Las Vegas, NV, USA), pp. 422–425, IEEE, Mar. 2018.

[5] A. Doryab, D. K. Villalba, P. Chikersal, J. M. Dutcher, M. Tumminia, X. Liu, S. Cohen, K. Creswell, J. Mankoff, J. D. Creswell, and A. K. Dey, "Identifying Behavioral Phenotypes of Loneliness and Social Isolation with Passive Sensing:

Statistical Analysis, Data Mining and Machine Learning of Smartphone and Fitbit Data," *JMIR mHealth and uHealth*, vol. 7, p. e13209, July 2019.

[6] I. Nahum-Shani, S. N. Smith, B. J. Spring, L. M. Collins, K. Witkiewitz, A. Tewari, and S. A. Murphy, "Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support," *Annals of Behavioral Medicine*, Sept. 2016.

[7] M. C. Sezgin, B. Gunsel, and G. K. Kurt, "Perceptual audio features for emotion detection," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2012, p. 16, May 2012.

[8] S. Saeb, M. Zhang, C. J. Karr, S. M. Schueller, M. E. Corden, K. P. Kording, and D. C. Mohr, "Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study," *Journal of Medical Internet Research*, vol. 17, p. e175, July 2015.

[9] S. Saeb, E. G. Lattie, K. P. Kording, and D. C. Mohr, "Mobile Phone Detection of Semantic Location and Its Relationship to Depression and Anxiety," *JMIR mHealth and uHealth*, vol. 5, p. e112, Aug. 2017.

[10] M. Boukhechba, Y. Huang, P. Chow, K. Fua, B. A. Teachman, and L. E. Barnes, "Monitoring social anxiety from mobility and communication patterns," in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers on - UbiComp '17*, (Maui, Hawaii), pp. 749–753, ACM Press, 2017.

[11] S. Mendu, A. Baglione, S. Baee, C. Wu, B. Ng, A. Shaked, G. Clore, M. Boukhechba, and L. Barnes, "A Framework for Understanding the Relationship between Social Media Discourse and Mental Health," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, pp. 144:1–144:23, Oct. 2020.

[12] J. A. Naslund, K. A. Aschbrenner, S. J. Kim, G. J. McHugo, J. Unützer, S. J. Bartels, and L. A. Marsch, "Health behavior models for informing digital technology interventions for individuals with mental illness.," *Psychiatric Rehabilitation Journal*, vol. 40, pp. 325–335, Sept. 2017.

[13] D. C. Klonoff, "Behavioral Theory: The Missing Ingredient for Digital Health Tools to Change Behavior and Increase Adherence," *Journal of Diabetes Science and Technology*, vol. 13, no. 2, pp. 276– 281, 2019.

[14] H. M. Jennings, J. Morrison, K. Akter, A. Kuddus, N. Ahmed, S. Kumer Shaha, T. Nahar, H. Haghparast-Bidgoli, A. A. Khan, K. Azad, and E. Fottrell, "Developing a theory-driven contextually relevant mHealth intervention," *Global Health Action*, vol. 12, p. 1550736, Jan. 2019. Publisher: Taylor & Francis.

[15] J. K. Salwen-Deremer, A. S. Khan, S. S. Martin, B. M. Holloway, and J. W. Coughlin, "Incorporating health behavior theory into mhealth: an examination of weight loss, dietary, and physical activity interventions," *Journal of Technology in Behavioral Science*, vol. 5, p. 51–60, Mar 2020.

[16] G. Eysenbach, "The Law of Attrition," *Journal of Medical Internet Research*, vol. 7, p. e11, Mar. 2005.

[17] P. Dourish, "What we talk about when we talk about context," *Personal and Ubiquitous Computing*, vol. 8, pp. 19–30, Feb. 2004.

[18] V. L. Champion and C. S. Skinner, "The health belief model.," in *Health behavior and health education: Theory, research, and practice, 4th ed.*, pp. 45–65, San Francisco, CA, US: Jossey-Bass, 2008.

[19] I. Ajzen, "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes*, vol. 50, pp. 179–211, 1991.

[20] D. Salber, A. Dey, and G. Abowd, "The context toolkit | Proceedings of the SIGCHI conference on Human Factors in Computing Systems," in *CHI '99: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, (Pittsburgh, Pennsylvania, USA), ACM, 1999.

[21] J. Torous, M. V. Kiang, J. Lorme, and J.-P. Onnela, "New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research," *JMIR Mental Health*, vol. 3, p. e16, May 2016.

[22] S. Shiffman, A. A. Stone, and M. R. Hufford, "Ecological Momentary Assessment," *Annual Review of Clinical Psychology*, vol. 4, pp. 1–32, Apr. 2008.

[23] W. T. Riley, D. E. Rivera, A. A. Atienza, W. Nilsen, S. M. Allison, and R. Mermelstein, "Health behavior models in the age of mobile interventions: are our

theories up to the task?," *Translational Behavioral Medicine*, vol. 1, pp. 53–71, Mar. 2011.

[24] D. C. Mohr, S. M. Schueller, E. Montague, M. N. Burns, and P. Rashidi, "The Behavioral Intervention Technology Model: An Integrated Conceptual and Technological Framework for eHealth and mHealth Interventions," *Journal of Medical Internet Research*, vol. 16, p. e3077, June 2014. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[25] H.-T. Cheng, S. Buthpitiya, F.-T. Sun, and M. L. Griss, "OmniSense: A Collaborative Sensing Framework for User Context Recognition Using Mobile Phones," in *The Eleventh International Workshop on Mobile Computing Systems and Applications*, (Annapolis, MD), p. 1, ACM, 2010.

[26] P. Chow, W. Bonelli, Y. Huang, K. Fua, B. A. Teachman, and L. E. Barnes, "DEMONS: an integrated framework for examining associations between physiology and self-reported affect tied to depressive symptoms," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16*, (Heidelberg, Germany), pp. 1139–1143, ACM Press, 2016.

[27] A. Doryab, P. Chikarsel, X. Liu, and A. K. Dey, "Extraction of Behavioral Features from Smartphone and Wearable Data," *arXiv:1812.10394 [cs, stat]*, Jan. 2019. arXiv: 1812.10394.

[28] N. Ramanathan, F. Alquaddoomi, H. Falaki, D. George, C.-K. Hsieh, J. Jenkins, C. Ketcham, B. Longstaff, J. Ooms, J. Selsky, H. Tangmunarunkit, and D. Estrin, "ohmage: An open Mobile System for Activity and Experience Sampling," in *Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare*, (San Diego, United States), IEEE, 2012.

[29] H. Tangmunarunkit, J. Kang, Z. Khalapyan, J. Ooms, N. Ramanathan, D. Estrin, C. K. Hsieh, B. Longstaff, S. Nolen, J. Jenkins, C. Ketcham, J. Selsky, F. Alquaddoomi, and D. George, "Ohmage: A General and Extensible End-to-End Participatory Sensing Platform," *ACM Transactions on Intelligent Systems and Technology*, vol. 6, pp. 1–21, Apr. 2015.

[30] D. Ferreira, V. Kostakos, and A. K. Dey, "AWARE: Mobile Context Instrumentation Framework," *Frontiers in ICT*, vol. 2, Apr. 2015.

[31] P. Chow, H. Xiong, K. Fua, W. Bonelli, B. A. Teachman, and L. E. Barnes, "SAD: Social Anxiety and Depression Monitoring System for College Students," p. 4, 2016.

[32] H. Xiong, Y. Huang, L. E. Barnes, and M. S. Gerber, "Sensus: a cross-platform, general-purpose system for mobile crowdsensing in human-subject studies," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16*, (Heidelberg, Germany), pp. 415–426, ACM Press, 2016.

[33] S. M. Hossain, T. Hnat, N. Saleheen, N. J. Nasrin, J. Noor, B.-J. Ho, T. Condie, M. Srivastava, and S. Kumar, "mCerebrum: a mobile sensing software platform for development and validation of digital biomarkers and interventions," in *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*, pp. 1–14, 2017.

[34] X. Liu, T. Chen, F. Qian, Z. Guo, F. X. Lin, X. Wang, and K. Chen, "Characterizing Smartwatch Usage In The Wild," in *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, (Niagara Falls, New York, USA), p. 14, Association for Computing Machinery, 2017.

[35] S. Intille, C. Haynes, D. Maniar, A. Ponnada, and J. Manjourides, " ema: Microinteraction-based ecological momentary assessment (ema) using a smartwatch.," *Proceedings of the ... ACM International Conference on Ubiquitous Computing . UbiComp (Conference)*, vol. 2016, p. 1124–1128, Sep 2016.

[36] C. Min, S. Kang, C. Yoo, J. Cha, S. Choi, Y. Oh, and J. Song, "Exploring current practices for battery use and management of smartwatches," in *Proceedings of the 2015 ACM International Symposium on Wearable Computers*, ISWC '15, (New York, NY, USA), p. 11–18, Association for Computing Machinery, 2015.

[37] S. Shiffman, A. A. Stone, and M. R. Hufford, "Ecological Momentary Assessment," *Annual Review of Clinical Psychology*, vol. 4, pp. 1–32, Apr. 2008.

[38] Y. Guo, Z. Xu, J. Qiao, Y. A. Hong, H. Zhang, C. Zeng, W. Cai, L. Li, and C. Liu, "Development and Feasibility Testing of an mHealth (Text Message and WeChat) Intervention to Improve the Medication Adherence and Quality of Life of People Living with HIV in China: Pilot Randomized Controlled Trial," *JMIR mHealth and uHealth*, vol. 6, p. e10274, Sept. 2018.

[39] D. Colombo, P. Cipresso, J. F. Alvarez, A. G. Palacios, G. Riva, and C. Botella, "An overview of factors associated with adherence and dropout to ecological momentary assessments in depression.," vol. 16, (US), p. 11–17, Interactive Media Institute, 2018.

[40] J. Torous, J. Lipschitz, M. Ng, and J. Firth, "Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis," *Journal of Affective Disorders*, vol. 263, pp. 413–419, Feb. 2020.

[41] J. Lehmann, M. Lalmas, E. Yom-Tov, and G. Dupret, "Models of User Engagement," in *User Modeling, Adaptation, and Personalization* (D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, J. Masthoff, B. Mobasher, M. C. Desmarais, and R. Nkambou, eds.), vol. 7379, pp. 164–175, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[42] A. Perianez, A. Saas, A. Guitart, and C. Magne, "Churn Prediction in Mobile Social Games: Towards a Complete Assessment Using Survival Ensembles," in

2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), (Montreal, QC, Canada), pp. 564–573, IEEE, Oct. 2016.

[43] S. Kim, D. Choi, E. Lee, and W. Rhee, "Churn prediction of mobile and online casual games using play log data," *PLOS ONE*, vol. 12, p. e0180735, July 2017.

[44] K. Cheung, W. Ling, C. J. Karr, K. Weingardt, S. M. Schueller, and D. C. Mohr, "Evaluation of a recommender app for apps for the treatment of depression and anxiety: an analysis of longitudinal user engagement," *Journal of the American Medical Informatics Association*, vol. 25, pp. 955–962, Aug. 2018.

[45] H.-M. Kim and B. K. Mallick, "A Bayesian prediction using the skew Gaussian distribution," *Journal of Statistical Planning and Inference*, vol. 120, pp. 85–101, Feb. 2004.

[46] J. Linardon and M. Fuller-Tyszkiewicz, "Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review.," *Journal of Consulting and Clinical Psychology*, vol. 88, pp. 1–13, Jan. 2020.

[47] J. Nicholas, J. Proudfoot, G. Parker, I. Gillis, R. Burckhardt, V. Manicavasagar, and M. Smith, "The Ins and Outs of an Online Bipolar Education Program: A Study of Program Attrition," *Journal of Medical Internet Research*, vol. 12, p. e57, Dec. 2010.

[48] O. Blanson Henkemans, W. Rogers, and A. Dumay, "Personal characteristics

and the law of attrition in randomized controlled trials of eHealth services for self-care," *Gerontechnology*, vol. 10, pp. 157–168, Dec. 2011.

[49] M. Price, D. F. Gros, J. L. McCauley, K. S. Gros, and K. J. Ruggiero, "Nonuse and Dropout Attrition for a Web-Based Mental Health Intervention Delivered in a Post-Disaster Context," *Psychiatry: Interpersonal and Biological Processes*, vol. 75, pp. 267–284, Sept. 2012. 00033.

[50] G. Kovacs, Z. Wu, and M. S. Bernstein, "Rotating Online Behavior Change Interventions Increases Effectiveness But Also Increases Attrition," *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, pp. 1–25, Nov. 2018.

[51] I. Chien, A. Enrique, J. Palacios, T. Regan, D. Keegan, D. Carter, S. Tschiatschek, A. Nori, A. Thieme, D. Richards, G. Doherty, and D. Belgrave, "A Machine Learning Approach to Understanding Patterns of Engagement With Internet-Delivered Mental Health Interventions," *JAMA Network Open*, vol. 3, pp. e2010791–e2010791, July 2020. _eprint: https://jamanetwork.com/journals/jamanetworkopen/articlepdf/2768347/chien_2020_oi_2

[52] X. Tang, Y. Liu, N. Shah, X. Shi, P. Mitra, and S. Wang, "Knowing your FATE: Friendship, Action and Temporal Explanations for User Engagement Prediction on Social Apps," *arXiv:2006.06427 [cs, stat]*, June 2020. arXiv: 2006.06427.

[53] F. Stojanovski, "Churn Prediction using Sequential Activity Patterns in an On-Demand Music Streaming Service," p. 83.

[54] C. Yang, X. Shi, L. Jie, and J. Han, "I Know You'll Be Back: Interpretable New User Clustering and Churn Prediction on a Mobile Social Application," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, (London United Kingdom), pp. 914–922, ACM, July 2018.

[55] J. Nicholas, J. Proudfoot, G. Parker, I. Gillis, R. Burckhardt, V. Manicavasagar, and M. Smith, "The Ins and Outs of an Online Bipolar Education Program: A Study of Program Attrition," *Journal of Medical Internet Research*, vol. 12, p. e57, Dec. 2010.

[56] E. Murray, I. R. White, M. Varagunam, C. Godfrey, Z. Khadjesari, and J. McCambridge, "Attrition Revisited: Adherence and Retention in a Web-Based Alcohol Trial," *Journal of Medical Internet Research*, vol. 15, p. e162, Aug. 2013.

[57] D. Ben-Zeev, E. A. Scherer, J. D. Gottlieb, A. J. Rotondi, M. F. Brunette, E. D. Achtyes, K. T. Mueser, S. Gingerich, C. J. Brenner, M. Begale, D. C. Mohr, N. Schooler, P. Marcy, D. G. Robinson, and J. M. Kane, "mHealth for Schizophrenia: Patient Engagement With a Mobile Phone Intervention Following Hospital Discharge," *JMIR Mental Health*, vol. 3, p. e34, July 2016.

[58] C. Blixen, M. Sajatovic, D. J. Moore, C. Depp, C. Cushman, J. Cage, M. Barboza, L. Eskew, P. Klein, and J. B. Levin, "Patient participation in the development of a customized m-Health intervention to improve medication adherence

in poorly adherent individuals with bipolar disorder (BD) and hypertension (HTN)," *International Journal of Healthcare*, vol. 4, p. 25, Mar. 2018.

[59] M. B. Schure, J. C. Lindow, J. H. Greist, P. A. Nakonezny, S. J. Bailey, W. L. Bryan, and M. J. Byerly, "Use of a Fully Automated Internet-Based Cognitive Behavior Therapy Intervention in a Community Population of Adults With Depression Symptoms: Randomized Controlled Trial," *Journal of medical Internet research*, vol. 21, no. 11, p. e14754, 2019. Publisher: JMIR Publications Inc., Toronto, Canada.

[60] H. Brath, J. Morak, T. Kästenbauer, R. Modre-Osprian, H. Strohner-Kästenbauer, M. Schwarz, W. Kort, and G. Schreier, "Mobile health (mHealth) based medication adherence measurement - a pilot trial using electronic blisters in diabetes patients: mHealth based medication adherence measurement system," *British Journal of Clinical Pharmacology*, vol. 76, pp. 47–55, Sept. 2013.

[61] N. Chiang, M. Guo, K. R. Amico, L. Atkins, and R. T. Lester, "Interactive Two-Way mHealth Interventions for Improving Medication Adherence: An Evaluation Using The Behaviour Change Wheel Framework," *JMIR mHealth and uHealth*, vol. 6, p. e87, Apr. 2018.

[62] M. E. Hilliard, A. Hahn, A. K. Ridge, M. N. Eakin, and K. A. Riekert, "User Preferences and Design Recommendations for an mHealth App to Promote Cystic Fibrosis Self-Management," *JMIR mHealth and uHealth*, vol. 2, p. e44, Oct. 2014.

[63] K. A. Kannisto, J. Korhonen, C. E. Adams, M. H. Koivunen, T. Vahlberg, and M. A. Välimäki, "Factors Associated With Dropout During Recruitment and Follow-Up Periods of a mHealth-Based Randomized Controlled Trial for Mobile.Net to Encourage Treatment Adherence for People With Serious Mental Health Problems," *Journal of Medical Internet Research*, vol. 19, p. e46, Feb. 2017.

[64] World Health Organization, "Adherence to Long-Term Therapies: Evidence for Action," 2003.

[65] C. C. Murphy, L. K. Bartholomew, M. Y. Carpentier, S. M. Bluethmann, and S. W. Vernon, "Adherence to adjuvant hormonal therapy among breast cancer survivors in clinical practice: a systematic review," *Breast Cancer Res Treat*, vol. 134, no. 2, pp. 459–78, 2012. Type: Journal Article.

[66] K. R. Jacob Arriola, T. A. Mason, K. A. Bannon, C. Holmes, C. L. Powell, K. Horne, and R. O'Regan, "Modifiable risk factors for adherence to adjuvant endocrine therapy among breast cancer patients," *Patient Education and Counseling*, vol. 95, pp. 98–103, Apr. 2014.

[67] C. Lin, R. Clark, P. Tu, H. B. Bosworth, and L. L. Zullig, "Breast cancer oral anti-cancer medication adherence: a systematic review of psychosocial motivators and barriers," *Breast Cancer Research and Treatment*, vol. 165, pp. 247–260, Sept. 2017.

[68] L. C. Markovitz, N. J. Drysdale, and B. A. Bettencourt, "The relationship between risk factors and medication adherence among breast

cancer survivors: What explanatory role might depression play?," *Psycho-Oncology*, vol. 26, no. 12, pp. 2294–2299, 2017. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pon.4362.

[69] L. M. Hess, A. Louder, K. Winfree, Y. E. Zhu, A. B. Oton, and R. Nair, "Factors Associated with Adherence to and Treatment Duration of Erlotinib Among Patients with Non-Small Cell Lung Cancer," *Journal of Managed Care & Specialty Pharmacy*, vol. 23, pp. 643–652, Mar. 2017. Publisher: Academy of Managed Care Pharmacy.

[70] P. M. Ho, C. L. Bryson, and J. S. Rumsfeld, "Medication adherence: its importance in cardiovascular outcomes.," *Circulation*, vol. 119, pp. 3028–3035, June 2009. Place: United States.

[71] M. El Alili, B. Vrijens, J. Demonceau, S. M. Evers, and M. Hiligsmann, "A scoping review of studies comparing the medication event monitoring system (MEMS) with alternative methods for measuring medication adherence: Scoping review of methods measuring medication adherence," *British Journal of Clinical Pharmacology*, vol. 82, pp. 268–279, July 2016.

[72] M. Aldeer, R. P. Martin, and R. E. Howard, "Pillsense: Designing a medication adherence monitoring system using pill bottle-mounted wireless sensors," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, 2018.

[73] T. Toscos, M. Drouin, J. A. Pater, M. Flanagan, S. Wagner, A. Coupe,

R. Ahmed, and M. J. Mirro, "Medication adherence for atrial fibrillation patients: triangulating measures from a smart pill bottle, e-prescribing software, and patient communication through the electronic health record," *JAMIA Open*, vol. 3, pp. 233–242, July 2020.

[74] H. Kalantarian, B. Motamed, N. Alshurafa, and M. Sarrafzadeh, "A wearable sensor system for medication adherence prediction," *Artificial Intelligence in Medicine*, vol. 69, pp. 43–52, May 2016.

[75] K. Stawarz, A. L. Cox, and A. Blandford, "Don't forget your pill! designing effective medication reminder apps that support users' daily routines," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, (New York, NY, USA), pp. 2269–2278, Association for Computing Machinery, Apr. 2014.

[76] D. Ben-Zeev, E. A. Scherer, J. D. Gottlieb, A. J. Rotondi, M. F. Brunette, E. D. Achtyes, K. T. Mueser, S. Gingerich, C. J. Brenner, M. Begale, D. C. Mohr, N. Schooler, P. Marcy, D. G. Robinson, and J. M. Kane, "mHealth for Schizophrenia: Patient Engagement With a Mobile Phone Intervention Following Hospital Discharge," *JMIR Mental Health*, vol. 3, p. e34, July 2016.

[77] M. Faurholt-Jepsen, M. Frost, E. M. Christensen, J. E. Bardram, M. Vinberg, and L. V. Kessing, "Validity and characteristics of patient-evaluated adherence to medication via smartphones in patients with bipolar disorder: exploratory reanalyses on pooled data from the monarca i and ii trials," *Evidence-Based Mental Health*, vol. 23, no. 1, pp. 2–7, 2020.

[78] H. Hafezi, T. L. Robertson, G. D. Moon, K. Au-Yeung, M. J. Zdeblick, and G. M. Savage, "An ingestible sensor for measuring medication adherence," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 1, pp. 99–109, 2015.

[79] B. A. McGregor, G. A. Vidal, S. A. Shah, J. D. Mitchell, and A. E. Hendifar, "Remote Oncology Care: Review of Current Technology and Future Directions," *Cureus*, vol. 12, no. 8.

[80] S. M. Cox, A. Lane, and S. L. Volchenboum, "Use of Wearable, Mobile, and Sensor Technology in Cancer Clinical Trials," *JCO Clinical Cancer Informatics*, pp. 1–11, Dec. 2018.

[81] M. Mohri, A. Rostamizadeh, and A. Talwalkar, "Regression," in *Foundations of Machine Learning*, pp. 267–294, MIT Press, 2 ed., 2012.

[82] C. Ikeda and D. Misaki, "Medication Monitoring from Accelerometer Data through a Series of Medication Actions Using Neural Network for Medication Adherence Evaluation," in *2022 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 288–291, Jan. 2022. ISSN: 2375-9356.

[83] G. Nguyen, S. Dlugolinsky, V. Tran, and A. Lopez Garcia, "Deep Learning for Proactive Network Monitoring and Security Protection," *IEEE Access*, vol. 8, pp. 19696–19716, 2020.

[84] V. Koesmahargyo, A. Abbas, L. Zhang, L. Guan, S. Feng, V. Yadav, and I. R. Galatzer-Levy, "Accuracy of machine learning-based prediction of medication

adherence in clinical research," *Psychiatry Research*, vol. 294, p. 113558, Dec. 2020.

[85] Y. Gu, A. Zalkikar, L. Kelly, K. Daly, and T. E. Ward, "Predicting Injectable Medication Adherence via a Smart Sharps Bin and Machine Learning," in *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, pp. 1–7, June 2020.

[86] Y. Gu, A. Zalkikar, M. Liu, L. Kelly, A. Hall, K. Daly, and T. Ward, "Predicting medication adherence using ensemble learning and deep learning models with large scale healthcare data," *Scientific Reports*, vol. 11, p. 18961, Dec. 2021.

[87] J. Cherian, S. Ray, and T. Hammond, "An Activity Recognition System for Taking Medicine Using In-The-Wild Data to Promote Medication Adherence," in *26th International Conference on Intelligent User Interfaces*, IUI '21, (New York, NY, USA), pp. 575–584, Association for Computing Machinery, Apr. 2021.

[88] G. S. Karanasiou, E. E. Tripoliti, T. G. Papadopoulos, F. G. Kalatzis, Y. Goletsis, K. K. Naka, A. Bechlioulis, A. Errachid, and D. I. Fotiadis, "Predicting adherence of patients with HF through machine learning techniques," *Healthcare Technology Letters*, vol. 3, pp. 165–170, Sept. 2016.

[89] X. Li, H. Xu, M. Li, and D. Zhao, "Using Machine Learning Models to Study Medication Adherence in Hypertensive Patients Based on National Stroke Screening Data," in *2021 IEEE 9th International Conference on Bioinformatics and Computational Biology (ICBCB)*, pp. 135–139, May 2021.

[90] R. Nieuwlaat, N. Wilczynski, T. Navarro, N. Hobson, R. Jeffery, A. Keepanasseril, T. Agoritsas, N. Mistry, A. Iorio, S. Jack, B. Sivaramalingam, E. Iserman, R. Mustafa, D. Jedraszewski, C. Cotoi, and R. Haynes, "Interventions for helping patients to follow prescriptions for medications," *Cochrane Database of Systematic Reviews*, no. 11, pp. 1–732, 2014.

[91] K. J. Wells, T. M. Pan, C. Vázquez-Otero, D. Ung, A. E. Ustjanauskas, D. Muñoz, C. Laronga, R. G. Roetzheim, M. Goldenstein, C. Carrizosa, S. Nuhaily, K. Johnson, M. Norton, E. Sims, and G. P. Quinn, "Barriers and facilitators to endocrine therapy adherence among underserved hormone-receptor-positive breast cancer survivors: a qualitative study," *Supportive Care in Cancer*, vol. 24, no. 10, pp. 4123–4130, 2016.

[92] V. S. Conn, T. M. Ruppar, M. Enriquez, and P. Cooper, "Medication adherence interventions that target subjects with adherence problems: Systematic review and meta-analysis," *Research in Social and Administrative Phramacy*, vol. 12, no. 2, pp. 218–246, 2016.

[93] K. Santo, S. S. Richtering, J. Chalmers, A. Thiagalingam, K. C. Chow, and J. Redfern, "Mobile Phone Apps to Improve Medication Adherence: A Systematic Stepwise Process to Identify High-Quality Apps," *JMIR Mhealth Uhealth*, vol. 4, p. e132, Dec. 2016.

[94] A. N. Baglione, J. Gong, M. Boukhechba, K. J. Wells, and L. E. Barnes, "Leveraging Mobile Sensing to Understand and Develop Intervention Strate-

gies to Improve Medication Adherence," *IEEE Pervasive Computing*, vol. 19, pp. 24–36, July 2020.

[95] L. Osterberg and T. Blaschke, "Adherence to Medication," *New England Journal of Medicine*, vol. 353, no. 5, pp. 487–497, 2005.

[96] A. Bandura, *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall series in social learning theory., Englewood Cliffs, NJ, US: Prentice-Hall, Inc, 1986. Type: Book.

[97] M. Boukhechba, S. Baee, A. L. Nobles, J. Gong, K. Wells, and L. E. Barnes, "A social cognitive theory-based framework for monitoring medication adherence applied to endocrine therapy in breast cancer survivors," in *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, (Las Vegas, NV, USA), pp. 275–278, IEEE, Mar. 2018.

[98] W. T. Riley, A. Cesar, Martin, and D. E. Rivera, "The importance of behavior theory in control system modeling of physical activity sensor data," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6880–6883, 2014. ISSN: 1094-687X.

[99] J. Gong, Y. Huang, P. I. Chow, K. Fua, M. S. Gerber, B. A. Teachman, and L. E. Barnes, "Understanding behavioral dynamics of social anxiety among college students through smartphone sensors," *Information Fusion*, vol. 49, pp. 57–68, Sept. 2019.

[100] L. R. Martin, S. L. Williams, K. B. Haskard, and M. R. DiMatteo, "The

challenge of patient adherence," *Therapeutics and Clinical Risk Management*, vol. 1, no. 3, pp. 189–199, 2005.

[101] D. Ben-Zeev, E. A. Scherer, R. Wang, H. Xie, and A. T. Campbell, "Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health.," *Psychiatric Rehabilitation Journal*, vol. 38, no. 3, pp. 218–226, 2015.

[102] M. Boukhechba, P. Chow, K. Fua, B. A. Teachman, and L. E. Barnes, "Predicting Social Anxiety From Global Positioning System Traces of College Students: Feasibility Study," *JMIR Mental Health*, vol. 5, p. e10101, July 2018.

[103] H. E. Payne, C. Lister, J. H. West, and J. M. Bernhardt, "Behavioral Functionality of Mobile Apps in Health Interventions: A Systematic Review of the Literature," *JMIR mHealth and uHealth*, vol. 3, p. e20, Feb. 2015.

[104] L. Bartholomew Eldredge, C. Markham, R. Ruiter, M. Fernández, G. Kok, and G. Parcel, *Planning health promotion programs: an Intervention Mapping approach*. United States: Jossey-Bass Inc., 4 ed., 2016.

[105] A. N. Baglione, S. Tolman, C. Dapaah, D. Johnson, K. J. Wells, R. Hall, R. Gentzler, and L. E. Barnes, "mHealth for Medication and Side Effect Monitoring: Patients' Attitudes Toward Smart Devices for Managing Oral Chemotherapy during Lung Cancer Treatment." 2022.

[106] American Cancer Society, "Lung cancer statistics: How common is lung cancer?," 2022.

[107] D. M. Geynisman and K. E. Wickersham, "Adherence to targeted oral anti-cancer medications.," *Discovery medicine*, vol. 15, pp. 231–241, Apr. 2013.

[108] L. Wood, "A review on adherence management in patients on oral cancer therapies," *European Journal of Oncology Nursing*, vol. 16, pp. 432–438, Sept. 2012.

[109] K. D. Burney, K. Krishnan, M. T. Ruffin, D. Zhang, and D. E. Brenner, "Adherence to single daily dose of aspirin in a chemoprevention trial: An evaluation of self-report and microelectronic monitoring," *Archives of family medicine*, vol. 5, no. 5, p. 297, 1996. Publisher: JAMA.

[110] C. Caldeira, P. Bhowmick, P. Komarlingam, and K. A. Siek, "A State-Based Medication Routine Framework," in *CHI Conference on Human Factors in Computing Systems*, (New Orleans LA USA), pp. 1–16, ACM, Apr. 2022.

[111] J. Pater, S. Owens, S. Farmer, E. Mynatt, and B. Fain, "Addressing medication adherence technology needs in an aging population," in *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, PervasiveHealth '17, (New York, NY, USA), pp. 58–67, Association for Computing Machinery, May 2017.

[112] J. Mauro, K. B. Mathews, and E. S. Sredzinski, "Effect of a Smart Pill Bottle and Pharmacist Intervention on Medication Adherence in Patients with Multiple Myeloma New to Lenalidomide Therapy," *Journal of Managed Care & Specialty Pharmacy*, vol. 25, pp. 1244–1254, Oct. 2019. Publisher: Academy of Managed Care Pharmacy.

[113] M. L. Jacobs, J. Clawson, and E. D. Mynatt, "My journey compass: a preliminary investigation of a mobile tool for cancer patients," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, (New York, NY, USA), pp. 663–672, Association for Computing Machinery, Apr. 2014.

[114] A. Pereira-Salgado, J. A. Westwood, L. Russell, A. Ugalde, B. Ortlepp, J. F. Seymour, P. Butow, L. Cavedon, K. Ong, S. Aranda, S. Breen, S. Kirsa, A. Dunlevie, and P. Schofield, "Mobile Health Intervention to Increase Oral Cancer Therapy Adherence in Patients With Chronic Myeloid Leukemia (The REMIND System): Clinical Feasibility and Acceptability Assessment," *JMIR mHealth and uHealth*, vol. 5, p. e8349, Dec. 2017. Company: JMIR mHealth and uHealth Distributor: JMIR mHealth and uHealth Institution: JMIR mHealth and uHealth Label: JMIR mHealth and uHealth Publisher: JMIR Publications Inc., Toronto, Canada.

[115] O. Schmalz, C. Jacob, J. Ammann, B. Liss, S. Iivanainen, M. Kammermann, J. Koivunen, A. Klein, and R. A. Popescu, "Digital Monitoring and Management of Patients With Advanced or Metastatic Non-Small Cell Lung Cancer Treated With Cancer Immunotherapy and Its Impact on Quality of Clinical Care: Interview and Survey Study Among Health Care Professionals and Patients," *Journal of Medical Internet Research*, vol. 22, p. e18655, Dec. 2020. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research La-

bel: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[116] C. A. Low, "Harnessing consumer smartphone and wearable sensors for clinical cancer research," *npj Digital Medicine*, vol. 3, pp. 1–7, Oct. 2020. Number: 1 Publisher: Nature Publishing Group.

[117] B. A. McGregor, G. A. Vidal, S. A. Shah, J. D. Mitchell, and A. E. Hendifar, "Remote Oncology Care: Review of Current Technology and Future Directions," *Cureus*, vol. 12, no. 8, 2020.

[118] O. Ciani, M. Cucciniello, F. Petracca, G. Apolone, G. Merlini, S. Novello, P. Pedrazzoli, N. Zilembo, C. Broglia, E. Capelletto, M. Garassino, E. Nicod, and R. Tarricone, "Lung Cancer App (LuCApp) study protocol: a randomised controlled trial to evaluate a mobile supportive care app for patients with metastatic lung cancer," *BMJ Open*, vol. 9, p. e025483, Feb. 2019. Publisher: British Medical Journal Publishing Group Section: Oncology.

[119] N. E. Billingy, V. N. M. F. Tromp, E. Veldhuijzen, J. Belderbos, N. K. Aaronson, E. Feldman, R. J. A. Hoek, H.-J. Bogaard, B. D. Onwuteaka-Philipsen, L. van de Poll-Franse, J. G. Hugtenburg, C. J. G. van den Hurk, A. Becker-Commissaris, and I. Walraven, "SYMptom monitoring with Patient-Reported Outcomes using a web application among patients with Lung cancer in the Netherlands (SYMPRO-Lung): study protocol for a stepped-wedge randomised controlled trial," *BMJ Open*, vol. 11, p. e052494, Sept. 2021.

[120] X. Skrabal Ross, K. M. Gunn, P. Patterson, and I. Olver, "Development of a Smartphone Program to Support Adherence to Oral Chemotherapy in People with Cancer," *Patient preference and adherence*, vol. 13, pp. 2207–2215, Dec. 2019.

[121] K. A. Kessel, M. M. Vogel, C. Kessel, H. Bier, T. Biedermann, H. Friess, P. Herschbach, R. von Eisenhart-Rothe, B. Meyer, M. Kiechle, U. Keller, C. Peschel, R. M. Schmid, and S. E. Combs, "Mobile Health in Oncology: A Patient Survey About App-Assisted Cancer Care," *JMIR Mhealth Uhealth*, vol. 5, p. e81, June 2017.

[122] N. D. Panayi, M. M. Mars, and R. Burd, "The promise of digital (mobile) health in cancer prevention and treatment.," *Future oncology (London, England)*, vol. 9, pp. 613–617, May 2013. Place: England.

[123] H. I. Portability and A. A. of 1996.

[124] H. I. T. for Economic and C. H. A. of 2009.

[125] H. L. . International, "Health level 7 standards," 2022.

[126] K. Theodos and S. Sittig, "Health Information Privacy Laws in the Digital Age: HIPAA Doesn't Apply.," *Perspectives in health information management*, vol. 18, no. Winter, p. 1l, 2021.

[127] X. Ni, Y. Lou, W. Hu, H. Wang, H. Xu, S. Li, Y. Zhou, and Y. Ni, "Development of mobile health–based self-management support for patients with lung cancer: A stepwise approach," *Nursing Open*, vol. 9, pp. 1612–1624, May 2022.

[128] S. Harper, J. Lynch, S. C. Meersman, N. Breen, W. W. Davis, and M. E. Reichman, "An Overview of Methods for Monitoring Social Disparities in Cancer with an Example Using Trends in Lung Cancer Incidence by Area-Socioeconomic Position and Race-Ethnicity, 1992-2004," *American Journal of Epidemiology*, vol. 167, pp. 889–899, Feb. 2008.

[129] W. T. Riley, C. A. Martin, D. E. Rivera, E. B. Hekler, M. A. Adams, M. P. Buman, M. Pavel, and A. C. King, "Development of a dynamic computational model of social cognitive theory," *Translational Behavioral Medicine*, vol. 6, pp. 483–495, Dec. 2016.

[130] C. A. Martín, D. E. Rivera, E. B. Hekler, W. T. Riley, M. P. Buman, M. A. Adams, and A. B. Magann, "Development of a Control-Oriented Model of Social Cognitive Theory for Optimized mHealth Behavioral Interventions," *IEEE Transactions on Control Systems Technology*, vol. 28, pp. 331–346, Mar. 2020. Conference Name: IEEE Transactions on Control Systems Technology.

[131] D. C. Mohr, K. N. Tomasino, E. G. Lattie, H. L. Palac, M. J. Kwasny, K. Weingardt, C. J. Karr, S. M. Kaiser, R. C. Rossom, L. R. Bardsley, L. Caccamo, C. Stiles-Shields, and S. M. Schueller, "IntelliCare: An Eclectic, Skills-Based App Suite for the Treatment of Depression and Anxiety," *Journal of Medical Internet Research*, vol. 19, p. e10, Jan. 2017.

[132] S. Abdullah and T. Choudhury, "Sensing Technologies for Monitoring Serious Mental Illnesses," *IEEE MultiMedia*, vol. 25, pp. 61–75, Jan. 2018.

[133] Q. Pham, G. Graham, C. Carrion, P. P. Morita, E. Seto, J. N. Stinson, and J. A. Cafazzo, "A Library of Analytic Indicators to Evaluate Effective Engagement with Consumer mHealth Apps for Chronic Conditions: Scoping Review," *JMIR mHealth and uHealth*, vol. 7, p. e11941, Jan. 2019.

[134] A. N. Baglione, L. Cai, A. Bahrini, I. Posey, M. Boukhechba, and P. I. Chow, "Understanding the Relationship Between Mood Symptoms and Mobile App Engagement Among Patients With Breast Cancer Using Machine Learning: Case Study," *JMIR Medical Informatics*, vol. 10, p. e30712, June 2022.

[135] S. García, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining*, vol. 72 of *Intelligent Systems Reference Library*. Cham: Springer International Publishing, 2015.

[136] S. García, S. Ramírez-Gallego, J. Luengo, J. M. Benítez, and F. Herrera, "Big data preprocessing: methods and prospects," *Big Data Analytics*, vol. 1, p. 9, Nov. 2016.

[137] A. Gupta, G. Ocker, and P. I. Chow, "Recruiting breast cancer patients for mHealth research: Obstacles to clinic-based recruitment for a mobile phone app intervention study," *Clinical Trials*, vol. 17, pp. 675–683, Dec. 2020. Publisher: SAGE Publications.

[138] J. E. Van Gemert-Pijnen, S. M. Kelders, and E. T. Bohlmeijer, "Understanding the Usage of Content in a Mental Health Intervention for Depression: An Analysis of Log Data," *Journal of Medical Internet Research*, vol. 16, p. e2991, Jan.

2014. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[139] T. Choudhary, L. N. Sharma, M. K. Bhuyan, and K. Bora, "Identification of Human Breathing-States Using Cardiac-Vibrational Signal for m-Health Applications," *IEEE Sensors Journal*, vol. 21, pp. 3463–3470, Feb. 2021. Conference Name: IEEE Sensors Journal.

[140] M. Boukhechba, A. R. Daros, K. Fua, P. I. Chow, B. A. Teachman, and L. E. Barnes, "DemonicSalmon: Monitoring Mental Health and Social Interactions of College Students Using Smartphones," *Smart Health*, July 2018.

[141] C. F. Bublitz, A. C. Ribeiro-Teixeira, T. A. Pianoschi, J. Rochol, and C. B. Both, "Unsupervised Segmentation and Classification of Snoring Events for Mobile Health," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pp. 1–6, Dec. 2017.

[142] J. A. Greer, J. M. Jacobs, N. Pensak, L. E. Nisotel, J. N. Fishbein, J. J. MacDonald, M. E. Ream, E. A. Walsh, J. Buzaglo, A. Muzikansky, I. T. Lennes, S. A. Safren, W. F. Pirl, and J. S. Temel, "Randomized Trial of a Smartphone Mobile App to Improve Symptoms and Adherence to Oral Therapy for Cancer," *Journal of the National Comprehensive Cancer Network*, vol. 18, no. 2, p. 11, 2020.

[143] T. Berger, E. Hohl, and F. Caspar, "Internet-based treatment for social phobia: a randomized controlled trial," *Journal of Clinical Psychology*, vol. 65, pp. 1021–1035, Oct. 2009.

[144] M.-H. Savard, J. Savard, A. Caplette-Gingras, H. Ivers, and C. Bastien, "Relationship between objectively recorded hot flashes and sleep disturbances among breast cancer patients: investigating hot flash characteristics other than frequency," *Menopause*, vol. 20, pp. 997–1005, Oct. 2013.

[145] F. Kronenberg, "Menopausal Hot Flashes: A Review of Physiology and Biosociocultural Perspective on Methods of Assessment," *The Journal of Nutrition*, vol. 140, pp. 1380S–1385S, July 2010.

[146] J. S. Carpenter, "The Hot Flash Related Daily Interference Scale: A Tool for Assessing the Impact of Hot Flashes on Quality of Life Following Breast Cancer," *Journal of Pain and Symptom Management*, vol. 22, pp. 979–989, Dec. 2001.

[147] S. Clarke, L. G. Jaimes, and M. A. Labrador, "mStress: A mobile recommender system for just-in-time interventions for stress," in *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, (Las Vegas, NV), pp. 1–5, IEEE, Jan. 2017.

[148] C.-H. Cho, T. Lee, M.-G. Kim, H. P. In, L. Kim, and H.-J. Lee, "Mood Prediction of Patients With Mood Disorders by Machine Learning Using Passive Digital Phenotypes Based on the Circadian Rhythm: Prospective Observational

Cohort Study," *Journal of Medical Internet Research*, vol. 21, p. e11029, Apr. 2019. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[149] J. Ma, A. Ovalle, and D. M.-k. Woodbridge, "Medhere: A Smartwatch-based Medication Adherence Monitoring System using Machine Learning and Distributed Computing," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4945–4948, July 2018. ISSN: 1558-4615.

[150] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '14 Adjunct*, (Seattle, Washington), pp. 3–14, ACM Press, 2014.

[151] BreastCancer.Org, "Breast cancer facts and statistics," Feb 2021.

[152] Centers for Disease Control and Prevention, "Breast Cancer Statistics," 2021.

[153] K. E. Weaver, L. P. Forsythe, B. B. Reeve, C. M. Alfano, J. L. Rodriguez, S. A. Sabatino, N. A. Hawkins, and J. H. Rowland, "Mental and Physical Health–Related Quality of Life among U.S. Cancer Survivors: Population Estimates

from the 2010 National Health Interview Survey," *Cancer Epidemiology and Prevention Biomarkers*, vol. 21, pp. 2108–2117, Nov. 2012. Publisher: American Association for Cancer Research Section: Research Articles.

[154] W. Linden, A. Vodermaier, R. Mackenzie, and D. Greig, "Anxiety and depression after cancer diagnosis: prevalence rates by cancer type, gender, and age," *Journal of Affective Disorders*, vol. 141, pp. 343–351, Dec. 2012.

[155] M. Reich, A. Lesur, and C. Perdrizet-Chevallier, "Depression, quality of life and breast cancer: a review of the literature," *Breast Cancer Research and Treatment*, vol. 110, pp. 9–17, July 2008.

[156] M. Watson, J. S. Haviland, S. Greer, J. Davidson, and J. M. Bliss, "Influence of psychological response on survival in breast cancer: a population-based cohort study," *Lancet (London, England)*, vol. 354, pp. 1331–1336, Oct. 1999.

[157] L. E. Carlson and B. D. Bultz, "Efficacy and medical cost offset of psychosocial interventions in cancer care: making the case for economic analyses," *Psycho-Oncology*, vol. 13, pp. 837–849; discussion 850–856, Dec. 2004.

[158] L. M. Gudenkauf, M. H. Antoni, J. M. Stagl, S. C. Lechner, D. R. Jutagir, L. C. Bouchard, B. B. Blomberg, S. Glück, R. P. Derhagopian, G. L. Giron, E. Avisar, M. A. Torres-Salichs, and C. S. Carver, "Brief cognitive-behavioral and relaxation training interventions for breast cancer: A randomized controlled trial," *Journal of Consulting and Clinical Psychology*, vol. 83, pp. 677–688, Aug. 2015.

[159] J. A. Johnson, J. A. Rash, T. S. Campbell, J. Savard, P. R. Gehrman, M. Perlis, L. E. Carlson, and S. N. Garland, "A systematic review and meta-analysis of randomized controlled trials of cognitive behavior therapy for insomnia (CBT-I) in cancer survivors," *Sleep Medicine Reviews*, vol. 27, pp. 20–28, June 2016. Publisher: W.B. Saunders Ltd.

[160] M. Chi, "The Hidden Cost of Cancer: Helping Clients Cope with Financial Toxicity," *Clinical Social Work Journal*, vol. 47, pp. 249–257, Sept. 2019.

[161] K. R. Yabroff, W. W. Davis, E. B. Lamont, A. Fahey, M. Topor, M. L. Brown, and J. L. Warren, "Patient time costs associated with cancer care," *Journal of the National Cancer Institute*, vol. 99, pp. 14–23, Jan. 2007.

[162] J. C. Holland, B. J. Kelly, and M. I. Weinberger, "Why psychosocial care is difficult to integrate into routine cancer care: stigma is the elephant in the room," *Journal of the National Comprehensive Cancer Network: JNCCN*, vol. 8, pp. 362–366, Apr. 2010.

[163] M. Charlton, J. Schlichting, C. Chioreso, M. Ward, and P. Vikas, "Challenges of Rural Cancer Care in the United States," *Oncology (Williston Park, N.Y.)*, vol. 29, pp. 633–640, Sept. 2015.

[164] O. Martínez Arroyo, Y. Andreu Vaíllo, P. Martínez López, and M. J. Galdón Garrido, "Emotional distress and unmet supportive care needs in survivors of breast cancer beyond the end of primary treatment," *Supportive Care in Cancer: Official Journal of the Multinational Association of Supportive Care in Cancer*, vol. 27, pp. 1049–1057, Mar. 2019.

[165] R. F. Muñoz, D. A. Chavira, J. A. Himle, K. Koerner, J. Muroff, J. Reynolds, R. D. Rose, J. I. Ruzek, B. A. Teachman, and S. M. Schueller, "Digital apothecaries: a vision for making health care interventions accessible worldwide," *mHealth*, vol. 4, June 2018.

[166] J. Firth, J. Torous, J. Nicholas, R. Carney, A. Pratap, S. Rosenbaum, and J. Sarris, "The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials," *World psychiatry: official journal of the World Psychiatric Association (WPA)*, vol. 16, pp. 287–298, Oct. 2017.

[167] G. Andersson, P. Cuijpers, P. Carlbring, H. Riper, and E. Hedman, "Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis," *World psychiatry: official journal of the World Psychiatric Association (WPA)*, vol. 13, pp. 288–295, Oct. 2014.

[168] M. A. Reger and G. A. Gahm, "A meta-analysis of the effects of internet- and computer-based cognitive-behavioral treatments for anxiety," *Journal of Clinical Psychology*, vol. 65, pp. 53–75, Jan. 2009.

[169] M. Sucala, P. Cuijpers, F. Muench, R. Cardoș, R. Soflau, A. Dobrean, P. Achimas-Cadariu, and D. David, "Anxiety: There is an app for that. A systematic review of anxiety apps: SUCALA et al.," *Depression and Anxiety*, vol. 34, pp. 518–525, June 2017.

[170] A. R. Wasil, K. E. Venturo-Conerly, R. M. Shingleton, and J. R. Weisz, "A review of popular smartphone apps for depression and anxiety: Assessing the inclusion of evidence-based content," *Behaviour Research and Therapy*, vol. 123, p. 103498, Dec. 2019.

[171] E. Mattila, R. Lappalainen, P. Välkkynen, E. Sairanen, P. Lappalainen, L. Karhunen, K. Peuhkuri, R. Korpela, M. Kolehmainen, and M. Ermes, "Usage and Dose Response of a Mobile Acceptance and Commitment Therapy App: Secondary Analysis of the Intervention Arm of a Randomized Controlled Trial," *JMIR mHealth and uHealth*, vol. 4, July 2016.

[172] R. Zhang, J. Nicholas, A. A. Knapp, A. K. Graham, E. Gray, M. J. Kwasny, M. Reddy, and D. C. Mohr, "Clinically Meaningful Use of Mental Health Apps and its Effects on Depression: Mixed Methods Study," *Journal of Medical Internet Research*, vol. 21, p. e15644, Dec. 2019.

[173] P. I. Chow, S. L. Showalter, M. Gerber, E. M. Kennedy, D. Brenin, D. C. Mohr, E. G. Lattie, A. Gupta, G. Ocker, and W. F. Cohn, "Use of Mental Health Apps by Patients With Breast Cancer in the United States: Pilot Pre-Post Study," *JMIR Cancer*, vol. 6, p. e16476, Apr. 2020.

[174] K. Kroenke, R. L. Spitzer, J. B. W. Williams, and B. Löwe, "An ultra-brief screening scale for anxiety and depression: the PHQ-4," *Psychosomatics*, vol. 50, pp. 613–621, Dec. 2009.

[175] B. M. Craig, B. B. Reeve, P. M. Brown, D. Cella, R. D. Hays, J. Lipscomb,

A. Simon Pickard, and D. A. Revicki, "US valuation of health outcomes measured using the PROMIS-29.," *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, vol. 17, pp. 846–853, Dec. 2014.

[176] S. Van Buuren and K. Groothuis-Oudshoorn, "mice: Multivariate Imputation by Chained Equations in R," *Journal of Statistical Software*, vol. 45, pp. 1–67, Dec. 2011. Number: 1.

[177] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, June 2002.

[178] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, (New York, NY, USA), pp. 785–794, Association for Computing Machinery, Aug. 2016.

[179] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds.), vol. 30, Curran Associates, Inc., 2017.

[180] C. Burgess, V. Cornelius, S. Love, J. Graham, M. Richards, and A. Ramirez, "Depression and anxiety in women with early breast cancer: five year observational cohort study," *BMJ (Clinical research ed.)*, vol. 330, p. 702, Mar. 2005.

[181] B. Grabsch, D. M. Clarke, A. Love, D. P. McKenzie, R. D. Snyder, S. Bloch, G. Smith, and D. W. Kissane, "Psychological morbidity and quality of life in women with advanced breast cancer: a cross-sectional survey," *Palliative & Supportive Care*, vol. 4, pp. 47–56, Mar. 2006.

[182] B. A. Vogel, J. Bengel, and A. W. Helmes, "Information and decision making: Patients' needs and experiences in the course of breast cancer treatment," *Patient Education and Counseling*, vol. 71, pp. 79–85, Apr. 2008.

[183] J. D. Harrison, J. M. Young, M. A. Price, P. N. Butow, and M. J. Solomon, "What are the unmet supportive care needs of people with cancer? A systematic review," *Supportive Care in Cancer: Official Journal of the Multinational Association of Supportive Care in Cancer*, vol. 17, pp. 1117–1128, Aug. 2009.

[184] C. G. Vivar and A. McQueen, "Informational and emotional needs of long-term survivors of breast cancer," *Journal of Advanced Nursing*, vol. 51, no. 5, pp. 520–528, 2005. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2648.2005.03524.x.

[185] A. R. Siebenhüner, M. Mikolasek, C. M. Witt, and J. Barth, "Improvements in Health Might Contradict Adherence to Mobile Health Interventions: Findings from a Self-Care Cancer App Study.," *Journal of alternative and complementary medicine (New York, N.Y.)*, vol. 27, pp. S115–S123, Mar. 2021. Place: United States.

[186] J. D. Elhai, J. C. Levine, R. D. Dvorak, and B. J. Hall, "Non-social features of smartphone use are most related to depression, anxiety and problematic

smartphone use," *Computers in Human Behavior*, vol. 69, pp. 75–82, Apr. 2017.

[187] W. Gao, H. Liu, C. Ge, X. Liu, H. Jia, H. Wu, and X. Peng, "A Clinical Prediction Model of Medication Adherence in Hypertensive Patients in a Chinese Community Hospital in Beijing," *American Journal of Hypertension*, vol. 33, pp. 1038–1046, Nov. 2020.

[188] S. Pinho, M. Cruz, F. Ferreira, A. Ramalho, and R. Sampaio, "Improving medication adherence in hypertensive patients: A scoping review," *Preventive Medicine*, vol. 146, p. 106467, May 2021.

[189] S. M. S. Islam, V. Mishra, M. U. Siddiqui, J. C. Moses, S. Adibi, L. Nguyen, and N. Wickramasinghe, "Smartphone Apps for Diabetes Medication Adherence: Systematic Review," *JMIR Diabetes*, vol. 7, p. e33264, June 2022. Company: JMIR Diabetes Distributor: JMIR Diabetes Institution: JMIR Diabetes Label: JMIR Diabetes Publisher: JMIR Publications Inc., Toronto, Canada.

[190] W.-H. Lo-Ciganic, J. M. Donohue, J. M. Thorpe, S. Perera, C. T. Thorpe, Z. A. Marcum, and W. F. Gellad, "Using Machine Learning to Examine Medication Adherence Thresholds and Risk of Hospitalization," *Medical care*, vol. 53, pp. 720–728, Aug. 2015.

[191] T. M. Deist, F. J. W. M. Dankers, G. Valdes, R. Wijsman, I.-C. Hsu, C. Oberije, T. Lustberg, J. v. Soest, F. Hoebers, A. Jochems, I. E. Naqa, L. Wee, O. Morin, D. R. Raleigh, W. Bots, J. H. Kaanders, J. Belderbos, M. Kwint, T. Solberg, R. Monshouwer, J. Bussink,

A. Dekker, and P. Lambin, "Machine learning algorithms for outcome prediction in (chemo)radiotherapy: An empirical comparison of classifiers," *Medical Physics*, vol. 45, no. 7, pp. 3449–3459, 2018. _eprint: https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1002/mp.12967.

[192] J. M. Franklin, C. Gopalakrishnan, A. A. Krumme, K. Singh, J. R. Rogers, J. Kimura, C. McKay, N. E. McElwee, and N. K. Choudhry, "The relative benefits of claims and electronic health record data for predicting medication adherence trajectory," *American Heart Journal*, vol. 197, pp. 153–162, Mar. 2018.

[193] M. Boukhechba and L. E. Barnes, "SWear: Sensing Using WEARables. Generalized Human Crowdsensing on Smartwatches," in *Advances in Usability, User Experience, Wearable and Assistive Technology* (T. Ahram and C. Falcão, eds.), Advances in Intelligent Systems and Computing, (Cham), pp. 510–516, Springer International Publishing, 2020.

[194] N. Kaur, M. Gonzales, C. G. Alcaraz, L. E. Barnes, K. J. Wells, and J. Gong, "Theory-Guided Randomized Neural Networks for Decoding Medication-Taking Behavior," in *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, (Athens, Greece), pp. 1–4, IEEE, July 2021.

[195] C. Esteban, O. Staeck, Y. Yang, and V. Tresp, "Predicting Clinical Events by Combining Static and Dynamic Information Using Recurrent Neural Networks," *arXiv:1602.02685 [cs]*, Nov. 2016. arXiv: 1602.02685.

[196] A. Leontjeva and I. Kuzovkin, "Combining Static and Dynamic Features for Multivariate Sequence Classification," in *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, (Montreal, QC, Canada), pp. 21–30, IEEE, Oct. 2016.

[197] M. H. Rahman, S. Yuan, C. Xie, and Z. Sha, "Predicting human design decisions with deep recurrent neural network combining static and dynamic data," *Design Science*, vol. 6, p. e15, 2020.

[198] J. O. Prochaska, S. Johnson, and P. Lee, "The Transtheoretical Model of behavior change.," in *The handbook of health behavior change, 3rd ed.*, pp. 59–83, New York, NY, US: Springer Publishing Company, 2009.

[199] S. P. Goldstein, B. C. Evans, D. Flack, A. Juarascio, S. Manasse, F. Zhang, and E. M. Forman, "Return of the JITAI: Applying a Just-in-Time Adaptive Intervention Framework to the Development of m-Health Solutions for Addictive Behaviors," *International Journal of Behavioral Medicine*, vol. 24, pp. 673–682, Oct. 2017.

[200] E. T. Hébert, E. M. Stevens, S. G. Frank, D. E. Kendzor, D. W. Wetter, M. J. Zvolensky, J. D. Buckner, and M. S. Businelle, "An ecological momentary intervention for smoking cessation: The associations of just-in-time, tailored messages with lapse risk factors," *Addictive Behaviors*, vol. 78, pp. 30–35, Mar. 2018.

[201] L. G. Jaimes, M. Llofriu, and A. Raij, "A stress-free life: just-in-time interventions for stress via real-time forecasting and intervention adaptation," in

Proceedings of the 9th International Conference on Body Area Networks, BodyNets '14, (Brussels, BEL), pp. 197–203, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Sept. 2014.

[202] M. Sharmin, A. Raij, D. Epstien, I. Nahum-Shani, J. G. Beck, S. Vhaduri, K. Preston, and S. Kumar, "Visualization of time-series sensor data to inform the design of just-in-time adaptive stress interventions," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15*, (Osaka, Japan), pp. 505–516, ACM Press, 2015.

[203] W. Hardeman, J. Houghton, K. Lane, A. Jones, and F. Naughton, "A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 16, p. 31, Apr. 2019.

[204] J. A. Pellowski, S. C. Kalichman, D. White, C. M. Amaral, G. Hoyt, and M. O. Kalichman, "Real-time Medication Adherence Monitoring Intervention: Test of Concept in People Living With HIV Infection," *Journal of the Association of Nurses in AIDS Care*, vol. 25, pp. 646–651, Nov. 2014.

[205] J. Schroeder, C. Wilkes, K. Rowan, A. Toledo, A. Paradiso, M. Czerwinski, G. Mark, and M. M. Linehan, "Pocket Skills: A Conversational Mobile Web App To Support Dialectical Behavioral Therapy," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, (Montreal QC, Canada), pp. 1–15, ACM Press, 2018.