NORTHWESTERN UNIVERSITY

Understanding Self-Tracked Data from Bounded Situational Contexts

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Technology and Social Behavior

By

Ada Ng

EVANSTON, ILLINOIS

September 2021

ABSTRACT

Over the past decade as smartphones and wearable tracking devices have grown in popularity, more individuals have begun collecting their own health and behavioral data. Innovations in sensor technology now allow individuals to continuously collect data over long periods of time with minimal effort. As a result, more data has become accessible for individuals and their healthcare providers to analyze and inform decisions. While these data are often assumed to be a record of health and a reflection of the self-tracker's routine living, it is inevitable that data is captured during a period of disruption or non-routine circumstances. This dissertation research investigated how self-tracked data that has been captured during such circumstances are reflected upon after the disruption has ended. The preliminary study explored the use of wearable data captured during an intensive outpatient therapy program, a circumstance outside of the patients' routine environment. The main study explored how women reflected on data from their pregnancy around one year after giving birth. Through these studies, I formulated the notion of bounded situational contexts to encapsulate how individuals perceive the boundaries of disruption within data that is captured. I discuss how self-tracking tools can be designed to enable individuals to modify visualizations of their data according to their perceived bounded situational contexts to aid in data interpretation.

ACKNOWLEDGEMENTS

The completion of this dissertation was only possible with the help of a broad network of friends and mentors. I am fortunate to know many people who were kind enough to lend their wisdom and show me compassion starting from long before I ever had the idea of earning a PhD. Along the way, I met others who were able to provide me guidance and companionship.

I am thankful for my advisor Madhu Reddy as he had welcomed me into the academic community and helped me navigate the world of conferences, publications, and all the reviews that were written and received in between. Through countless drafts and revisions, he has molded me into the academic I am today.

I am also grateful to have such supportive committee members, Stephen Schueller, Nabil Alshurafa, and Maia Jacobs who lit the spark that started the collaborations that resulted in these projects that I present in this document. Thank you for sharing your expertise with me.

I am indebted to my family for taking on the incalculable cost of allowing me the time and distance to obtain a higher education. As is the common tale of immigration, Mom and Dad sacrificed their own education to provide their children with opportunities in an unfamiliar country. Being away from home for so long is hard; luckily, I have Anne, Alberta, Kevin, and Alan holding down the fort for all these years. I am privileged to have your support.

It’s impossible to fully convey how important it is to find personal champions during a graduate program. Ethan Robison, Ashley Walker, Jordan Eschler, Siddharth Muthukumaran: you spend your time lifting up others from the pits of anxiety and dread through positivity and just by

being so darn knowledgeable. Rawan Alharbi, Zachary Gibson, Angela Smith, Renwen Zhang: I am constantly in awe of your brilliance and your future students are incredibly lucky.

Graduate school is also an experience that cannot be complete without community. PITCH Lab, TSB and MTS faculty, students, and staff provided me the foundation upon which I could succeed. The Northwestern Grad Gardens group nourished me with fresh air and vegetables. I am deeply grateful for my friends. We worked together in-person and remotely for months; we celebrated each other's wins and commiserated in our losses. Elizabeth Danielson, Nicole Tilly, Phoebe Jean-Pierre: you're all champs.

In the spirit of this dissertation, I want to also highlight the super stars in my life before and after graduate school. Lee Stamper and Nora Mastny: thank you for the encouragement through this life transition and for making Chicago a home. Sanghee Oh and Su Hyun Lee: thank you for the reminder that there has always been and always will be life outside these walls.

LIST OF ABBREVIATIONS

BSC bounded situational context

EMA ecological momentary assessment

HCI human-computer interaction

ITP intensive treatment program

PGD patient-generated data

PTSD post-traumatic stress disorder

RQ research question

sPGD sensor-based patient-generated data

I dedicate this dissertation to my family, the Cropsey Crew.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Motivation

The mass appeal of quick self-optimization through the datafication of the self has led to an explosion in popularity of wearable self-tracking devices over the past decade [203]. With every upgrade, leaders of the wrist-worn wearable industry, Fitbit and the Apple Watch, race to add features that will quantify a new aspect of the wearer's body and package the data back to the wearer in a visually appealing format. Beyond the mainstream channels, other wearables companies have developed products used to track oxygen levels in newborns[1], performance in professional athletes[2], local pollution[3], and fertility[4]. Users of wearables span the range from the soldiers[5] to pets[6]. At its foundation of this narrative of self-optimization is the ability to measure the self and reflect on data.

Prior studies have demonstrated how self-tracking enables individuals to better understand and improve their health by shedding light onto previously invisible habits and behaviors [142]. For instance, by using health-monitoring applications, individuals with diabetes can find correlations between their actions and their health which then can lead to changes in self-management practices [157]. In another example, providing individuals access to pedometers has been shown to increase their physical activity [5]. These studies suggest that tracking can

---

[1] https://owletcare.com/
[2] https://www.catapultsports.com/
[3] https://www.laroche-posay.me/en/article/MY-SKIN-TRACK-UV/a37392.aspx
[4] https://www.avawomen.com/
[5] https://www.nytimes.com/2018/01/29/world/middleeast/strava-heat-map.htmlp
[6] https://www.whistle.com/

motivate behavior change through the identification of direct causes and effects [110]. In recent years, researchers, designers, and developers have begun to explore the use of ubiquitous sensing technologies to detect and predict periods of stress, depression, schizophrenia, among other mental health conditions [50, 64, 65, 255].

In the clinic, self-tracking symptoms or activities have long been recommended to patients as a part of the treatment process [30, 125]. Research has suggested tracking items such as physical activity can help patients understand their depression or tracking moments a patient relies on alcohol in order to identify triggers [73, 163]. Sharing self-tracked data can supplement and enrich existing clinical health records in ways that were previously impossible or infeasible. Self-tracking technologies have been shown to open up opportunities for remote monitoring and more data-driven and collaborative decision-making [39]. As evidenced by the increasing inclusion of self-tracked data into electronic medical records, these data which were captured outside of the clinical environment are playing an increasingly important role in the health care system.

The commercialization of ubiquitous sensing technology has introduced complications for health data interpretation: these patient data are no longer captured solely within a controlled clinical environment. Patient health data are now captured often using sensors throughout different contexts of the individual’s life [186]. When activities and behaviors are captured by sensors, an individual’s actions are transformed into digital representations and key contextual information is often lost. There is a growing body of work within the HCI and ubiquitous computing fields around the development of systems that capture supplementary contextual information. These studies which investigated context-aware technology and the incorporation of context into self-tracked data often focus on specific aspects of context that can be captured passively such as geographical, temporal, or social contexts [73, 143, 208]. The contextual information intended to enrich the behavior captured are often represented in data visualizations at a fine-grained scale (i.e. on a moment-by-moment basis), for example, an individual's GPS location at discrete points in time [208]. There is often a gap in the information necessary to fully realize the tracker's situation as coarse-grained contexts (i.e. situations that span across individual moments) are often excluded in data visualizations.

1.1.1 Research Motivation: Non-Routine Circumstances

The few studies that have focused on recontextualizing self-tracked data beyond the fine-grained scale have primarily focused on routine circumstances within chronic physical health. I consider routine circumstances as a type of coarse-grained context because the concept of *routine* spans across individual moments. For instance, Raj et al. [193] investigated how the temporal context of "weekend" influenced blood sugar spikes in a regular pattern. Therefore, as opposed to focusing on fine-grained context – individual blood sugar readings at specific points in time – they studied how cyclical patterns across multiple *routine* weekends influenced blood sugar. However, not all of life is routine. Few studies on self-tracking focus on non-routine circumstances, yet these periods are critical to take into consideration when interpreting health data.

Breaks in routine circumstances are inevitable. Reasons for a routine to be disrupted or a transition to a new routine include moving homes, having a child, or dissolving a marriage [89, 99]. Yet, this information is often absent from data visualizations that individuals use to manage their health. For instance, an individual experiencing a depressive episode may fail to track their health during this period. Without visibility into this circumstance, viewers of the data may interpret lapses in tracking as simply forgetfulness when in fact, the missing data could indicate the need for an intervention [160]. Data representations that do not reflect non-routine circumstances may lead to misinterpretation and missed opportunities to gain an understanding about the self-tracker.

The process of interpreting data from a non-routine circumstance may be different from the process of interpreting data from routine circumstances. During routines, the individual is at a steady state and within a predictable context. Context can be comprised of both external and internal components [178]. External contexts can be physical constraints (e.g. environments, tools) or social influences (e.g. cultural norms, obligations) while internal contexts can be personal goals, skills, and preferences [106, 178]. For instance, a routine "weekend" can involve a set of recurring and predictable external contexts and internal motivations. During a non-routine circumstance, external contexts may be unfamiliar to the individual and therefore *internal context* plays a more critical role in sensemaking and subsequent activities [178, 254]. In this research, I explored how individuals reflect on self-tracked data captured during non-routine periods of life.

Through a series of studies, I develop the concept of a *bounded situational context* (BSC) as a sensitizing concept to refer to a personalized context that an individual implicitly defines by factoring internal and external contexts during non-routine circumstances. This concept helps to frame the observed temporal misalignment between the external disruption as well as the data that individuals internally perceived to be related to the disruption. Through this characterization, I aim to support the design of self-tracking technologies that more appropriately convey the lived experiences of the tracker. In turn, this will enable individuals to reflect on data they find relevant to their health and inform their health self-management practices.

1.2 Research Objectives

In this research, I examined the use of self-tracked data towards an understanding of mental health during non-routine circumstances. Motivated by a gap in prior research on the applications of passively captured self-tracked data in mental health clinical care settings, this dissertation first examined how mental health care providers envision using self-tracked data during a three-week inpatient treatment program, a non-routine circumstance. Results from this preliminary study prompted me to continue to explore how another non-routine circumstance, pregnancy, can influence how self-tracked data is used towards an understanding of health by individuals themselves.

This research has three objectives: (a) examine how mental health care providers envision the use of self-tracked data in clinical care, (b) understand how self-trackers envision using data from a non-routine circumstance and (c) identify design considerations when visualizing self-tracked data captured during a non-routine circumstance. Overall, I aim to expand research on self-tracking by focusing on mental health and providing insight on the care provider perspective and then by homing in on how non-routine circumstances affect how the resultant data is reflected upon. Through this research, I address the following questions (see Table 1):

Table 1: Summary of research questions, gaps, and contributions

Research Gap	Research Questions (#)	Contributions
Given the increase in use of sensor data by people to self-track their behaviors and activities, we need a better understanding of how self-tracked data can be used for mental health care from the perspective of mental health care providers.	(1) How can self-tracked data be used in mental health clinical care? *(preliminary study)* (1a) What are mental health care providers' perceptions of self-tracked data? (1b) What are the benefits and challenges of using self-tracked data in mental health clinical care settings?	An empirical investigation on the perspective of mental health care providers that treat patients on the use of patient self-tracked data for mental health resulted in a descriptive contribution expanding the body of literature on how self-tracking technology can improve mental health within clinical practice.
Despite the importance of understanding context that underlies self-tracked data, there is a lack of research on how data from non-routine circumstances are interpreted and used over time (i.e. once a routine returns or a new routine is established).	(2a) How do self-trackers define which data captured during non-routine circumstances are useful to reflect upon? *(main study)* (2b) How do self-trackers envision reflecting on data after the non-routine period has passed?	By examining data interpretation practices within non-routine circumstances, this study extended current literature that had focused primarily on routine circumstances. Descriptive contributions included extending current literature on how individuals interpret and perceive health data captured for a non-routine context. Theoretically, this study was the first to characterize how self-trackers implicitly employ data work in order to segment data from non-routine periods – defining the term *bounded situational context.*
Prior research on visualizations of self-tracked	(3a) How should we think about designing data visualization-	The empirical study surfaced assumptions of current self-

data have explored how to include context on a routine basis (e.g. weekends vs weekday behaviors) or on a fine-grained scale (e.g. instances of physical activity). However, little is known about how to represent data captured during non-routine circumstances, a type of coarse-grained context.	modification features that support the reflection of data after the non-routine period has passed? *(main study)* (3b) What are the benefits and challenges of modifying data visualizations that depict a non-routine period?	tracking data visualizations that fail to consider disruptions to routine. Design contributions included data visualization-modification features that can be used to depict bounded situational contexts and empirically evaluating their perceived utility.

1.3 Study Overview

This dissertation research consists of (1) a preliminary study in which I conducted interviews to understand care providers' perceptions of using patient wearable data during outpatient mental health treatment and (2) a main study in which I employed interviews and design elicitation methods to (a) understand how women interpret stress data captured during their pregnancy and (b) develop prototypes of visual-modification features both as a probe and to evaluate the features themselves.

Preliminary study: I explored how care providers envisioned using self-tracked data in care for their patients, veterans undergoing intensive outpatient treatment for post-traumatic stress disorder. As part of this treatment, veterans were removed from their routine lives for three weeks as they received care and voluntarily tracked data on themselves using a commercial wrist-worn wearable. Using this tool, the veterans collected data during a non-routine circumstance: a treatment program purposefully away from their home and their routine.

I interviewed thirteen health care providers at a clinic which included therapists, non-therapists, and administrators. In this study, the participants described how they might use the veterans' tracked data in order to provide care. A few participants voiced concerns that the data tracked during treatment may not be reflective of their patients' health once they complete treatment and return to their home and work environment. In other words, participants questioned the relevance of data captured in non-routine circumstances when trying to assess patient health. This prompted me to pursue a deeper understanding of self-tracking data captured during non-routine circumstances for the Main study.

Main study (Phase 1): To investigate the self-tracker perspective, I conducted interviews with 8 women who tracked their stress for ~14 weeks of their pregnancy using a biosensor that measured heart rate and by answering ecological momentary assessments about stress. Through this research, I sought an understanding of how they perceived stress data collected from pregnancy, a non-routine circumstance. During interviews, most participants also discussed how they might use data they collected from periods before pregnancy that are not reflective of their current routine. During the session, the participants were asked to describe how they wish to explore these data and the perceived value of these data over time given that their circumstances have changed.

Main study (Phase 2): Following the first phase of the main study, I created graphs of different types of health data simulated to depict a pre-pregnancy period, pregnancy, and post-pregnancy period. I conducted 13 more interviews with other women that had self-tracked their health during their pregnancy and asked them how they may want to modify each graph such that the visualizations display information that is relevant to them.

1.4 Dissertation Overview

This dissertation consists of six chapters as follows:

- Chapter 1: Outlines research motivation, objectives, and study design
- Chapter 2: Provides an overview of relevant research in the areas of self-tracking in health management, the role of context in self-tracking, and designing self-tracked data representations
- Chapter 3: Describes the methodologies employed in this research: interviews, elicitation, and design evaluation
- Chapter 4: Presents the preliminary study, answers the first research question, and motivates the main studies
- Chapter 5: Presents the main study, which answers the remaining research questions
- Chapter 6: Summarizes the research contribution of the dissertation and outlines future works

2 RELATED LITERATURE

In this section, I describe the literature from the HCI and Health Informatics communities focusing on self-tracking, the role of context, and data visualizations. First, I provide an overview of prior work on the use of self-tracking in healthcare. Next, I introduce prior research on self-tracking specifically in the mental health domain. Then, I review studies that investigate how context influences the interpretation of self-tracked data. Finally, I discuss prior work that explore how to design data visualizations that aim to contextualize self-tracked data.

2.1 Self-Tracking Defined

For centuries, people have participated in the practice of collecting data on themselves. The terms self-tracking, self-monitoring, and life logging are commonly used to refer to the act of recording one's own behaviors, thoughts, and feelings [134]. Thoresen and Mahoney [234] outline self-monitoring as a two-step process: (1) identifying the occurrence of the behavior and (2) systematically recording this observation. In 2007, the term "quantified-selfers" emerged to describe a culture of digital self-tracking, and now refers to a community of avid self-trackers [35, 152]. These data captured by individuals, when introduced into a clinical environment has been referred to as patient-generated data (PGD) or patient-generated health data [216].

Advances in technology have led to changes in how self-tracked data are captured. Manual and passive methods of tracking each have their own benefits and drawbacks. Manual tracking affords the tracker freedom and flexibility to track according to their personalized needs [10].

However, it is also considered high burden because it requires the tracker to actively participate in their tracking. Therefore, forgetting to track can lead to missing data. While burdensome, manual tracking can promote behavior change and impart reflection by encouraging deliberate actions that bring mindfulness to the behavior that is being tracked [132]. More recently, there has been a shift from manual tracking to passive tracking (i.e. using sensors); the wearable tracking device industry is growing exponentially year after year [109]. Sensors enable lay-persons to systematically capture data that was previously virtually impossible to capture such as the number of steps taken in one day or how much UV light their skin has been exposed to over the course of the day [219]. However, in contrast to manual tracking, automated self-tracking systems are typically less personalizable to lay-persons and the low-burden of tracking also led to lower self-reflection [132]. In sum, different self-tracking methods have inherent trade-offs.

2.2 Opportunities and Challenges in Self-tracking for Health

A key aspect of self-tracking is that it renders activities and behaviors that would otherwise be invisible into visible, tangible assets. Self-tracked health data can offer benefits not only for self-management of health, but also for clinical care. However, there are multifaceted challenges to the realization of these benefits.

2.2.1 Data for the Individual

For individuals self-tracking of their own volition (i.e. not clinician-initiated), data can inform and incite action [35]. For example, people diagnosed with a chronic disease such as irritable bowel syndrome and diabetes have used self-monitoring to conduct n-of-1 self-experimentation activities in order to learn more about their health [121, 156, 181]. An individual with diabetes can create a personal blood glucose index by tracking their blood glucose after eating specific foods [156]. These insights in turn can lead to changes in eating behaviors. In these instances, tracking allowed individuals to access health data that can reveal patterns such as symptom triggers and the efficacy of interventions.

However, there are challenges to self-tracking including burdensome data collection and lack of tools to interpret the collected data [35]. Manual self-tracking is burdensome because it relies on individuals' ability to remember to track, to have the appropriate tools to track, and to perform the action of tracking [8, 43]. After the data is collected, some trackers find that there is no easy way to analyze and interpret data. This issue is particularly salient for trackers who are not skilled in statistics or data manipulation. While automated tracking circumvents the barrier of burdensome data collection, the large quantity of data collected exacerbates the issue of data interpretation [36]. Furthermore, commercial sensors are often accompanied with out-of-the-box data visualizations that provide visual summaries of data that are often designed in a one-size-fits-all format [118]. This can result in a misalignment between the self-tracker's goals and the data that is presented back to them. For example, a study on women with eating disorders found that food tracking applications further encouraged unhealthy calorie restriction [54]. The complexity of the dynamics of self-monitoring activities and the enabling technology is further compounded when these data are introduced into the clinical arena.

2.2.2 Data for Clinical Care

In contrast to traditional health data captured within the clinic, patient-generated data (PGD) are significant because they reflect patients' health, activities, and behaviors in their daily life. PGD can also provide a more complete record of the patient's health information that can be used to improve diagnoses and facilitate the personalization of treatment [38]. Furthermore, by actively monitoring their own health, patients are more empowered in their discussion about their health goals and priorities with their providers [39]. The availability of these data can also allow healthcare providers to monitor patients remotely and intervene in a timely matter [104]. In sum, the use of PGD in the clinic can improve clinician awareness of the patient's health as well as improve patient-provider collaboration [146].

While there are benefits to incorporating PGD into the clinical environment, adoption has been low due to a variety of factors [4, 38, 251, 259]. The value of PGD in applied settings is still unclear; clinicians aim to mitigate risk and harm when engaging with patients [22]. Therefore, they are often skeptical of the reliability of PGD because the data is collected in uncontrolled settings and devices are often not clinically validated [250]. There are also technical barriers to adoption. For instance, the lack of secure electronic medical record integration with PGD-capturing devices can restrict a clinician's ability to review the data [259]. In addition to these barriers, clinicians' limited time dedicated to each patient leads them to favor familiar and traditional clinical health measurements and practices of care when making decisions [250]. Consequently, integrating PGD into clinical care is unlikely to occur without technical, social, and organizational change [259].

2.3 Self-tracking in Mental Health

Mental well-being is related to positive social and physical health outcomes [7, 248]. However, most research on digital self-tracking for health management has focused on chronic physical health conditions. There is a gap in our understanding when it comes to its applications to mental health care [250].

While self-tracking of behaviors, moods, thoughts, and symptoms manually has a long history in clinical mental health care, the use of passively sensed data in mental health is a new and growing area of research [73, 95, 163, 230]. Digital self-tracking tools for mental health care commonly fall into two classes: (1) digitization of traditional tools of care and (2) new sources of data that complement or replace traditional tools. The first category consists of digitalization of previously paper tools such as questionnaires (e.g. MoodTracker [154], diaries [161], etc.). Given that these tools build upon existing practices, they may see similar benefits and challenges to clinical adoption as digital self-tracking tools for physical health conditions. For instance, clinicians may already be familiar with specific measures of health and therefore are confident in data interpretation as compared to new sources of health data. Additionally, lack of technical integration to electronic medical records lead to the same obstacles seen in the physical health space.

Tools that fall into the second class, those that leverage technology to replace or augment traditional clinical health data, have the potential to change how mental health issues can be detected, prevented, and treated. However, few studies have investigated how clinicians might apply these new tools in practice [162, 231]. Instead of depending on patient self-report which may be unreliable [220, 241], sensors can enable healthcare providers to directly monitor their patients for physiological and behavioral signals of problematic mental health episodes or adherence to therapeutic recommendations. However, while the use of sensors to monitor physical conditions such as diabetes can be straightforward, its applications to mental health conditions such as depression can be indirect and ambiguous. Indeed, within clinical mental health care, most research on the use of sensors focuses on increasing physical activity as a means of managing mental health issues [2, 37, 42, 207]. Furthermore, while there is a growing body of research on the use of machine learning to track mental health status [173, 231], few investigations have included the perspective of mental health clinicians. More research with mental health care providers themselves is necessary to identify more opportunities for the use of sensor data in mental health clinical care settings.

2.4 Context and Self-Tracking

Unlike traditional health data captured by a healthcare provider at point of care, self-tracked data collected by patients is affected by the different contexts of their daily lives. When patients capture data passively by wearing a sensor or actively by answering questions periodically, these data captured are inherently influenced by the various factors within the patient's ongoing life circumstances. For instance, blood glucose data from a patient, when collected at a clinician's office will reflect a specific and moderately controlled value while blood glucose data from a continuous blood glucose monitor, collected throughout the day, will reflect variations throughout the day or days of the week based on various life contexts such as work and leisure [193]. An understanding of these different contexts in which the health data is collected is imperative for accurate data interpretation. Below, I examine prior research in HCI on incorporating context when capturing or using self-tracked data. I identified a gap in the current literature on context and self-tracking; currently, routine and frequently recurring contexts have been the primary focus of study. By examining the few prior studies on self-tracking during non-routine contexts such as disruptions, I argue that disruptions are common. Therefore, an understanding of non-routine context is important for individuals and providers to be aware of when reflecting on self-tracked data.

2.4.1 Context Defined

> "*Context is any information that can be used to characterise the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.*" – Anind K. Dey (1999) [1]

Context is relevant information that, when absent, would change meaning of what is being communicated. Context is multidimensional (e.g. physical/spatial, social, emotional [187] and in a given point in time, different types of context can coexist, be nested within each other, and hierarchical [51, 189]. Understanding the multiple facets of contextual information can contribute to individuals' understanding of their health status and health management activities [34, 181].

Contextual information is a critical factor when it comes to the interpretation of self-tracked data because context is deeply embedded in each stage of tracking [59]. These stages include the motivation to track, the means of capturing data, the ongoing process of collecting, integrating, reflecting on the data, and potential lapses and resuming of tracking. For instance, with regards to motivations to track, contexts such as health condition and external circumstances such as social contexts can influence one's intention or ability to track. The information collected will reflect the envisioned uses and desired insights from the data. Regarding the means of capturing data, tracking through manual or automated approaches depend on technical and economic contexts. For example, the availability of affordable mobile technology and the individual's wealth can determine the feasibility of holding on to manual tracking tools during the day or wearing sensors. Ultimately, the data collected reflects the various dimensions of context across the different stages of self-tracking.

As more people begin to track their health using wearable sensors, long-term tracking through a variety of changing life contexts is becoming increasingly common, which introduces new factors to consider when interpreting data [166, 200]. Meyer et al. [166] recently defined long-term tracking as tracking that occurs over the span of years. When individuals track in the long-term, lapses in tracking are expected and understanding the contexts that have led to these lapses may be as meaningful as the data itself. Furthermore, when data persists in perpetuity, it is unclear how long records will be meaningful, useful, and relevant to the individual. For instance, heart rate data captured during stressful period of illness may no longer be relevant once the individual has made a full recovery. Therefore, capturing the contexts surrounding these data become increasingly pertinent when interpreting data that has been collected over long periods of time.

2.4.2 Routine Contexts

Context-aware systems often rely on dimensions of context that are detected on a moment-by-moment basis, such as a specific location at a specific time, which I refer to as fine-grained context. Pavel et al. [187] classified contexts that contribute to an understanding of how various aspects of life influences wellbeing into the following dimensions: physical/spatial, social, emotional, mental, activity, availability, and environmental. The dimensions of context are numerous and as such, much research on context-aware systems have focused on how systems can be designed to capture contextual information passively [188]. Breaking down context into discrete entities provides a foundation for researchers to collect variables through various passive data collection tools such as sensors and APIs. For instance, Rabbi et al. [191] used motion, light, temperature, humidity, and audio sensors to assess wearers' mental and physical health. Outside of automated data collection methods, researchers have also developed models that rely on capturing data through analog means such as ecological momentary assessments in which individuals periodically record their activity or mood such as stress in-the-moment [175]. Fine-grained contextualized systems enable important features such as just-in-time interventions [1, 114, 201, 217, 240] and user-friendly data aggregates such as averages and trends [18]. Indeed, aggregation is more readily understandable by individuals:

> "*One would expect our added contextual cues (such as location visits and commutes) to strengthen users' capacity to reconstruct past days, which should make past days' history more meaningful and interaction more likely. This was not supported…Such low-level representations of context have been found to be less valuable in uncovering the factors that influence behaviors when compared to high-level representations of physical activity (e.g., tables with overall exercise performed during a week at work), as they require 'paging days of detailed data to attempt to find trends, correlations, or opportunities for change.'*" [83]

In such instances, the more data that is collected, the more accurate predictions will be. However, the underlying assumption in these studies is that the data is collected during a routine period in which the data captured during these times are transferrable to future instances.

Routines are repeated actions that are caused by recurring situations [100, 101, 174]. Routines consist of: (1) a context and (2) the set of behaviors in response to that context. Therefore, routine contexts enable predictable outcomes [184]. For instance, the routine context of *weekend* is recurring and can be used to detect and predict measures of health related to weekend-related environments and behaviors [193]. Frequently recurring departures from routine are referred to as routine variations and in themselves, can be predictable, for example, monthly happy hours that lead to anticipated spikes in alcohol consumption [12]. However, it is common for routines to break and resume, or change entirely.

2.4.3 Non-Routine Contexts

Infrequent or non-recurring deviations from routine may lead to uncharacteristic actions that are less predictable, for instance, a natural disaster that leads to altered patterns of eating [100, 101]. Deviations from routine can be a time of instability or unpredictability and have been characterized in various ways in the HCI literature. Scholars have referred to the period after a destabilizing life disruption or life event such as dissolution, moving, or gender transition as a *life transition* or a construction of a *new normal* [89, 99, 159, 210]. Such unexpected "crises" and transitional periods that "punctuate discernable time blocks of a more or less steady state" (i.e. routine) can be unpredictable and difficult to define for patients, their caregivers, and healthcare providers [26]. The experience of cancer has been described as a journey, a period of life or health in flux [113, 167, 226]. Sandbulte et al. [206] refer to disruptive moments such as illness that change how family members share their health information as "turning points." **Broadly, I will refer to these periods of instability, transition, construction of a potential new normal, which are not inherently recurring, as *non-routine* contexts.**

While disruptions or changes to routine are a common occurrence, few studies have characterized the nature of self-tracking or the use of the resulting data during these non-recurring periods of disruption, transition, or evolution [89, 99, 200]. Prior HCI research that has explored self-tracking during such periods have primarily been situated in chronic illness contexts. For instance, prior work studying the lived experiences of self-tracking Parkinson's Disease has found that individuals and their caregivers experience cycles of *routinization* of health management tasks as well as evolving routines through disease progression [169, 170, 180, 242]. Outside of the context of chronic illness, as more data on individuals become available through long-term self-tracking, understanding how individuals interpret a lifetime of data – through its continuous evolution and the process of aging – will be increasingly relevant. Early work on self-reflection through long-term quantified tracking by Elsden et al. found that data were used to aid individuals in creating personal accounts of their past [56]. On a granular level, prior research on self-tracking for health have suggested that there are several ways in which non-routine contexts can influence both the tracked data itself as well as the interpretation of data that is tracked.

2.4.4 Interpreting Self-Tracked Data from Non-Routine Contexts

Given the challenges of ambiguity of and variety in types of disruptions, we know little about how self-tracked data collected during non-routine contexts are interpreted, how their relevance evolves across changing contexts, and how we should design data visualizations for such situations [182]. However, by examining prior studies specifically on lapses in the routine self-tracking, I found shifts in how self-tracked data are collected and interpreted during such contexts.

Lapses in self-tracking have been long studied in HCI and commonly framed as a 'failure' to track and a missed opportunity for insight [138, 166, 200]. However, by considering the context of a disruption to the tracker's individual lives, lapses in data can be reframed as valuable information. For instance, Matthews et al. [160] found that due to the "syncopated rhythms" of bipolar disorder, the onset of an episode of depressive symptoms may lead to a struggle to self-track. Thus, the absence of data may be clinically meaningful in itself. Schroder et al. [213] introduced the concept of *nonbinary lapsing* when they studied the self-tracking practices of individuals that experience migraines. While most literature up until that point had framed self-tracking as a binary of tracking or not tracking [138, 166, 200], they found that individuals would modify the type of information they recorded during stressful life circumstances as opposed to completely lapsing. These studies suggest that knowledge around the non-routine contexts that are not regularly recurring or inherently cyclical could serve as information to help future data-viewers interpret gaps or modifications in what is tracked as valuable insights about health. While disruptions can be reflected in data through gaps or changes in what is tracked, they can also manifest as values captured that deviate from the norm.

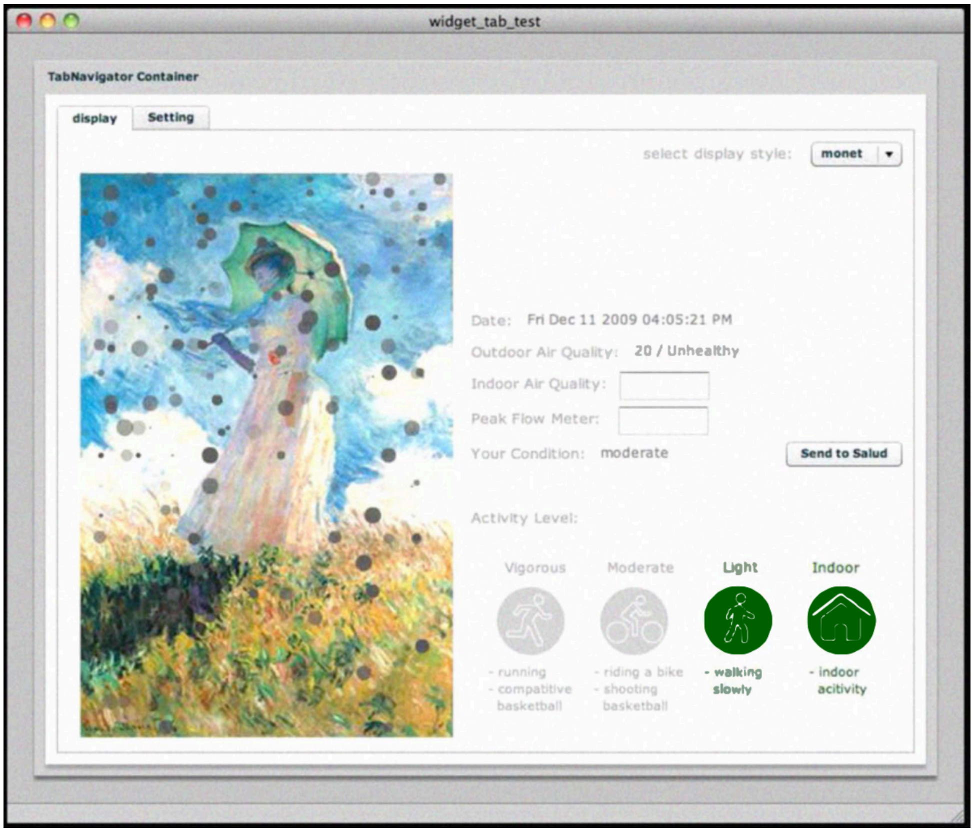
Circumstances of disruptions can influence the data that is captured and how they are perceived [200]. During the COVID-19 pandemic, individuals in quarantine had to adjust their exercise plans because goals including a high step count were no longer reasonably achievable [145, 227]. Body weight and food tracking data during pregnancy also deviate from that data of a prior routine (i.e. from a non-pregnant period). Without the context of "pandemic" or "pregnancy," the data collected during these disruptions (a low step count or rapid weight gain) may be considered alarming. Visibility into contexts of disruption or transition are arguably even more valuable when interpreting self-tracked data than during routine periods [143]. Given the importance of understanding context when interpreting self-tracked data, there is a growing body of research in HCI that investigates how to make contextual information available within visualizations of self-tracked data [181].

2.5 Designing for Recontextualization

Visualization is a powerful tool for making sense of large quantities of data. Visualized health data can quickly convey information about individuals' health condition and feedback regarding their health trajectory [181]. Interfaces for visualizing data may display a single value (e.g. blood glucose monitor) or multiple values such as mood over time (e.g. MONARCA) [13]. How data and their associated contexts are presented impact the understanding that individuals glean from reviewing these data.

2.5.1 Leveraging Context in Design

With the ever-increasing amount of health data being collected by sensors, context becomes more important to support interpretation. Context in various forms are used to enrich data; this may include potential explanatory factors or perhaps triggers for an individual's memory to promote self-reflection [181]. Contextual information has been presented in the form of photos, text, weather information, and phone activity among others. Context that has been incorporated into visualizations, as previously discussed, are often at a fine-grained scale. For instance, Figure 1 depicts an artistic representation of outdoor and indoor air quality with respect to a tracker's activities to allow them to reflect on how air quality impacts their asthma symptoms [256]. Such designs exclude other contextual factors may be co-occurring such as a period of illness or transition to new medication that may be affecting asthma symptoms. These factors are critical for an individual to be aware of when interpreting these data.



select display style: monet

Date: Fri Dec 11 2009 04:05:21 PM

Outdoor Air Quality: 20 / Unhealthy

Indoor Air Quality:

Peak Flow Meter:

Your Condition: moderate

Figure 1: Visualization of air quality and types of activities to help a tracker reflect on their asthma status [256].

2.5.2 Designing Visualizations of Context

Within the health data visualization and context space, there are studies that explore how individuals will interpret contextualized data and those that investigate visualization methods.

In an example of the former, Sharmin et al. [217] created various data visualization prototypes based on input from experts and evaluated each visualization with potential users. The goal of Sharmin et al.'s visualizations were to provide researchers with a visualized overview of their participants' stress data in order for them to identify appropriate times to present just-in-time stress interventions. Their designs were informed by a group of biomedical researchers, with whom they collaborated. The authors did not outline how they created their designs, but ultimately, they proposed four visualizations: spatio-temporal, temporal, contextual, and event-centric. Their results suggest that each of the visualization types provided different levels of granularity of context, each with its own advantages and disadvantages. The focus of such studies are to understand the use of contextualized data more so than designing such visualizations [87].

In an example of a study that focuses more on *how* to design visualizations for context, Epstein et al. [58] created visualizations of self-tracked data intended for personal use and self-reflection. First, they collected self-trackers' goals and interests through a survey. They then introduced the concept of *visual cuts* which feature a subset of collected data with specific geographic or physical activity data to contextualize the values displayed. Based on the surveys, they developed various combinations of contextual features using a variety of visualization methods including tables, graphs, captions, maps, and Sankey diagrams. The goal of this study was to assess the type of visualization method most appropriate to convey the desired information. Similar to the results of Sharmin et al.'s study, they found that the contextual information was a valuable tool for data interpretation.

In these examples, the contextual data undoubtedly provided additional insight about stress or other self-tracked data. However, the contexts depicted were on a fine-grained scale of individual days or instances of physical activity. Taking a broader view, when data exists in perpetuity, data collected from the past periods of life may not be representative of the individual in their present situation and therefore aggregating these data may obscure the current realities. Once a routine context has changed and a new normal has been established or once a disruption has passed and a routine context has resumed, how should such historical data be represented in visualizations?

2.5.3 Conceptualizing Context Through Design

In a literature review on self-care technologies, Nunes et al.'s [181] identified six design considerations to help designers determine how to provide contextualized information specifically with respect to chronic conditions. They suggest: (1) patients more knowledgeable about their condition may seek to reflect on multiple dimensions of context for a single data point, (2) a patient should focus on one specific value and explore the context around it, (3) overview visualizations that show how a specific value changes over time must include context in order to avoid superficial interpretation, (4) only contextual information that is capturable should be incorporated into visualizations, (5) it may be beneficial to combine both overview visualizations and contextualized visualizations and (6) multiple visualization perspectives should be used to engage different types of users. However, when considering the context of a disruption, we see that it is less clear about how data from disruptions should be presented based on these strategies. For instance, a disruption may occur at an unexpected time with an unknown duration. This raises the question whether the designers should be making the decisions around how to present this coarse-grained context or whether the individual themselves should control how coarse-grained contexts are represented in the visualization.

My thesis work is complementary to Raj et al.’s [193] study in which they introduce the notion of contextual frames: “*recurring contexts in which the factors that influence behavior and health status take on different degrees of relevance, thereby having different impacts on behavior.*” For instance, a contextual frame of *summer camp* can lead to a patient’s lack of access to appropriate foods and increased activity, which may lead to low blood glucose levels. Through understanding these patterns, researchers and designers can create representations of data that more closely embody the lived experiences that influence health-related behaviors. In their discussion of design implications when considering contextual frames, Raj et al. [193] note that (1) timescale representations of data must reflect the type of temporal cycles as determined by users, (2) certain contextual frames may have relationships with one another, possibly reflecting a sequential pattern which can then be used for health predictions, and (3) similarities and differences between contextual frames can be used to design effective support strategies. In contrast, the main study of this dissertation seeks to explore the uses for self-tracked data captured during disruptions and potentially non-recurring contexts. Given the lack of cyclical patterns, not all the findings about contextual frames apply to contexts of disruption. For instance, receiving a bone marrow transplant is a disruptive period of transition and uncertainty which introduces novel and unfamiliar factors that influence patients’ health [26]; such factors place the context of recovery outside of the definition of a contextual frame. This raises the question of how we design visualizations of self-tracked data captured during contexts that are newly emerging (i.e. not yet be classified as a contextual frame) or perhaps will never recur.

2.6 Summary

Designing features to surface distinct periods of non-routine life that are reflected in self-tracked data can support trackers in their goal of learning about or managing their health. To do this, we first need to better understand the uses for self-tracked data that is captured during non-routine contexts. I defined non-routine contexts broadly as periods of disruption, instability, transition, and construction of a potential new normal. I investigated how data collected during these periods can be used in the moment, proximally, and distally.

The HCI community is beginning to approach the design of personal informatics systems as less of a static tool and more as a dynamic and subjective class of tools that allows for changing goals and life experiences [111, 151, 185]. Prior HCI research on recontextualizing self-tracked data have primarily investigated routine circumstances such as weekends and workdays or routine variations within chronic physical illness and physical health. I acknowledge that current technology is not designed to capture non-routine contexts and therefore logging such contexts are likely to be manual and highly burdensome [143, 158]. However, if we can better understand how these types of data can be used or change how people perceive their data, we can then identify ways to design systems towards meaningful data collection, curation, and utilization.

3 METHODOLOGY

3.1 Chapter Overview

The preliminary and main research studies are formative and qualitative in nature. Formative research is often used when exploring new conceptualizations and generating knowledge in an emerging area [74]. Qualitative research methods take an interpretive approach and are used to understand another's view of the world rather than that of my own [116]. These methods are appropriate for both my preliminary study in which I sought to understand the perspective of mental health care providers and in the main study in which I investigated how women come to understand data from their pregnancy with respect to other periods of life. This chapter provides a brief overview of the data collection and analysis methods employed in this dissertation work. Details on the research methods for the preliminary and main studies are provided in Chapter 4 and Chapter 5 respectively.

3.2 Data Collection

Data was gathered through semi-structured interviews with text-based elicitation methods in the preliminary study and through semi-structured interviews with visual elicitation activities in the main study.

3.2.1 Semi-structured interviews

Semi-structured interviews are a commonly used method in HCI research and in particular, in healthcare research [79, 229]. The interview guide is versatile and flexible, allowing for spontaneity in response from participants. It also provides structure so that the researchers can collect similar types of information from each participant while allowing for participants' personal experiences to surface [103]. After allowing for participants' personal expressions, interviewers may improvise follow-up questions [202]. Therefore, the construction of the semi-structured interview guide must adequately allow the researcher to gain an understanding of the phenomenon studied, prevent the collection of data that is not necessary for the research, and empower participants to voice their individual experiences.

3.2.1.1 Generative Elicitation Activities

While semi-structured interviews traditionally involve the elicitation of information through questioning, elicitation techniques can be used to enrich question-based responses [16, 236]. Visual elicitation methods such as photo elicitation can be used to trigger memory in participants, lead to new perspectives, even help develop rapport with the participant [61, 117]. Oral-based elicitation such as, vignettes, stories about an individual or a situation, are presented to participants through text or orally before they are asked for their response [107, 250]. Vignettes can be used to investigate hypothetical responses or concepts unfamiliar to the participant. Broadly, semi-structured interviews with elicitation activities consist of prompts which provide structure to capture information towards the study purpose and the guiding research questions [119, 126].

My studies employ text-based and graphic elicitation methods in conjunction with semi-structured interview questions. In my preliminary study with mental health care providers, I used text-based elicitation cards as a means of reminding participants of the list of items to discuss including types of data captured by wearables that they may be unfamiliar with.

Similarly, my main study with pregnant or recently pregnant women involved selecting from an array of cards with types of self-tracked data on each according to prompts. These text-based elicitation methods have been employed by HCI researchers in the past [57, 90, 112, 171].

Additionally, for the main study I used visual elicitation techniques to complement semi-structured interview questions; specifically, I employed graphic elicitation. Graphic elicitation techniques such as diagrams have been used to depict conceptual subjects through graphic abstraction [45]. An example of a graphic elicitation activity used is the *timeline* (see Figure 2) [11]. By having participants draw a timeline of their past, they can reflect on their past and present lives; this is critical as a core component of the main research study is understanding the temporality of context when it comes to self-tracked data. Drawing activities allows participants to reflect on the topic being explored [77]. Diagrams such as the timeline can be used to provide a basis for further probing and communication between researcher and participant [45].

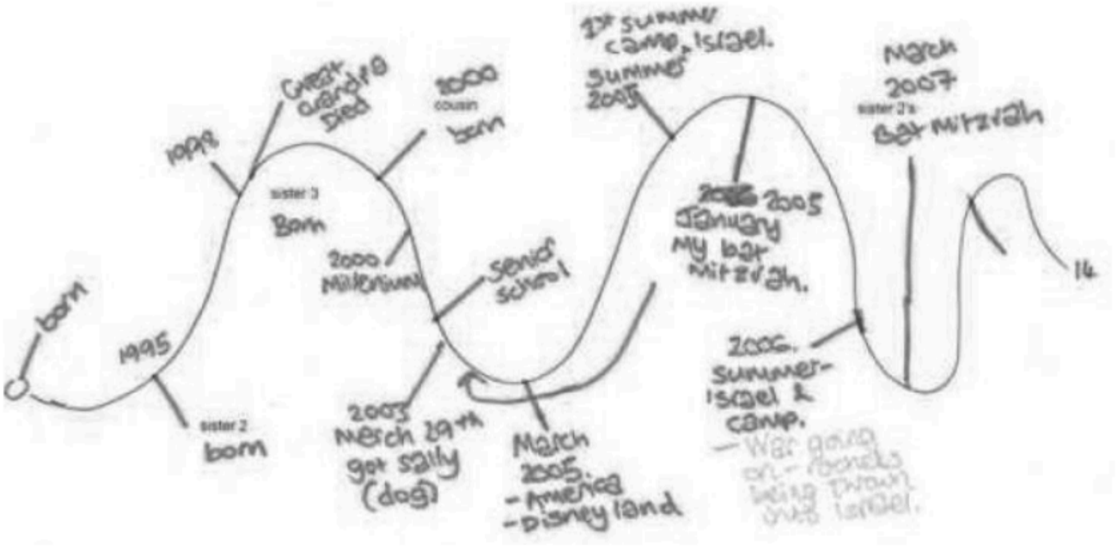


Figure 2. A participant-drawn timeline from [11]

3.2.1.2 Evaluative Elicitation Activities

Within HCI, researchers employ a variety of methods to evaluate whether a design has achieved the intended goal [195]. The choice for evaluation method is derived from the research goal [85]. For instance, a self-reported user questionnaire may be effective evaluating subjective measures that such as 'Satisfaction,' but a task analysis may be more appropriate when evaluating the efficiency that a design enables.

As one primary objective of the main study was to better understand how data should be visualized, graphic elicitation was a key component of the evaluation process [102]. Through the semi-structured interviews in Phase 1, participants expounded upon their envisioned uses (or non-uses) of their health data captured before, during, and after their pregnancy. Participants described how they perceived the bounded situational nature of their pregnancy data. Based on findings from these initial interviews, I developed different means of modifying data visualizations. In Phase 2 of the main study, participants evaluated the utility of these data modification features through a visual demonstration of a graph of simulated pregnancy data during a video call. My research methods follow the trajectory demonstrated in Crilly et al.'s model of the graphic elicitation process [45] (Figure 3).

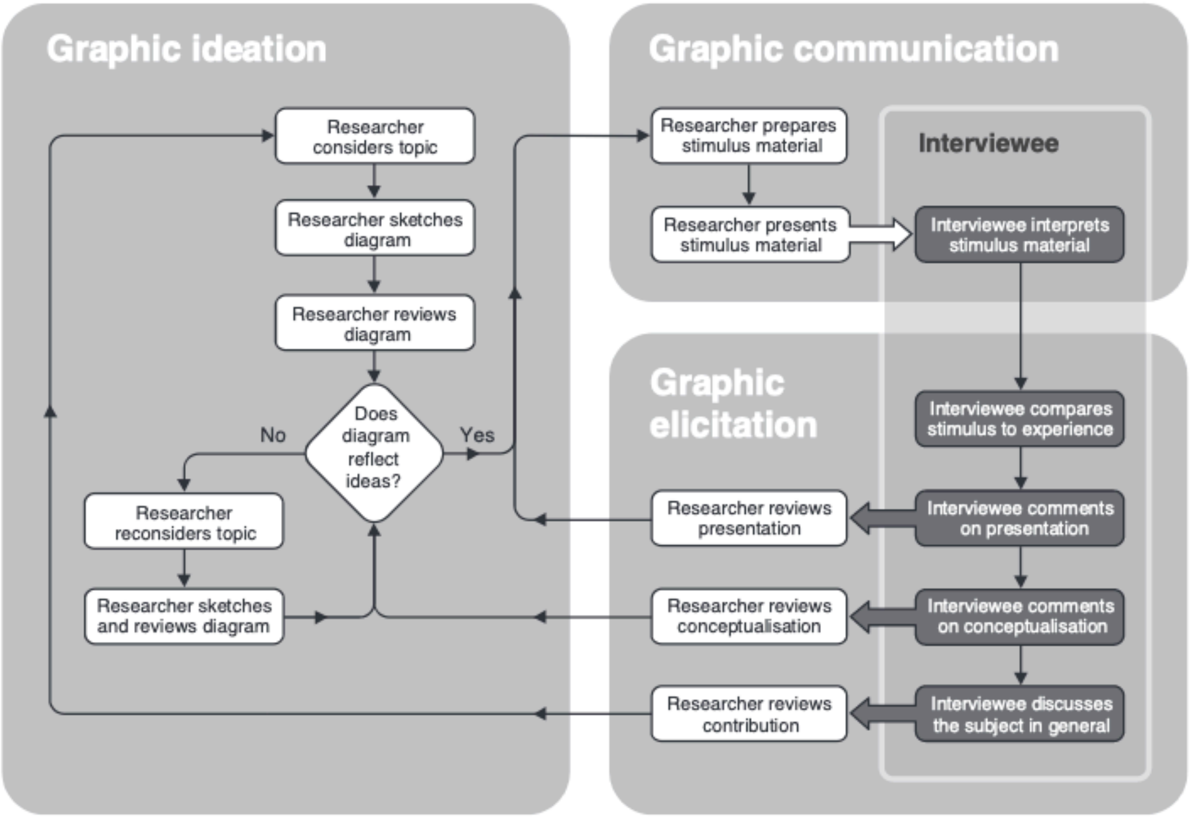


Figure 3. A flowchart of the graphic elicitation process from [45].

3.2.1.3 Interviewing Effectively

Once a semi-structured interview guide has been designed, it is the role of the researcher to conduct the interview effectively. Researcher skills include making the participant feel comfortable and safe to disclose information, clear communication, active listening, the ability to direct and redirect a conversation towards the research goals tactfully, and the ability to engage while maintaining neutrality. In my preliminary study, I conducted individual semi-structured interviews with mental health care providers. The semi-structured interview protocol allowed the participants to share their lived experiences that they may not have previously given much thought about and allowed me as the researcher to reflect their conceptualizations back to them as I guided the discussion. Similarly, in my main study, I explored how participants conceptualize data that they might not have previously thought of in-depth. By employing a semi-structured interview protocol with question-based prompts complemented with text- and other oral-based prompts, I was able to elicit participants' stories and lived experiences.

3.3 Data Analysis

In each of these studies, I aimed to convey the story of participants' perspectives. Therefore, I employed Braun and Clarke's thematic analysis in my data analysis [23].

3.3.1 Thematic Analysis

Throughout my data collection for all stages of this research, I used the thematic analysis method [23]. Within HCI, thematic analysis is a widely used tool for analyzing qualitative data. Using this method, themes are identified, analyzed, and patterns are interpreted from data. Braun and Clarke outlined this process in six phases (Table 2).

Table 2. Braun & Clarke's Six-Phase Thematic Analysis Procedure.

Phase		Description of the process
1	Familiarizing oneself with the data	Transcribing data; reading and re-reading the data; noting down initial codes
2	Generating initial codes	Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code
3	Searching for the themes	Collating codes into potential themes, gathering all data relevant to each potential theme
4	Involved reviewing the themes	Checking if the themes work in relation to the coded extracts and the entire data set; generating a thematic 'map'

5	Defining and naming themes	Ongoing analysis to refine the specifics of each theme; generating clear definitions and names for each theme
6	Producing the report	Final opportunity for selecting appropriate extracts; discussion of the analysis; relate back to research question or literature; produce report

Data analysis and data interpretation occurred during and after data collection [23]. Due to the visual and remote nature of the main study, videos were reviewed for visual patterns and since this activity is leveraged to prompt further probing and verbal communication, quotes were spot transcribed as part of the interview data corpus [96]. Throughout data collection, I reviewed audio or video recordings as appropriate while writing memos and formulating open codes inductively and iteratively. Once a comprehensive set of open codes were developed, I created axial codes, abstracting the open codes into a higher-level conceptualization. These axial codes were applied to other session videos and iterated upon to see where more prominent themes emerged, and which codes do not fit and therefore needed to be discarded. After iteratively defining these codes across the entire dataset, I defined and refined themes for the final analysis. These final themes guided the results and discussion of the study.

3.3.2 Validity and Transferability

Within HCI, qualitative research is often used to understand complex interactions between life and technology [3]. For at least a quarter of a century, there has been disagreement among researchers on what consisted “quality” qualitative research [19, 198]. However it is agreed that the goal of qualitative research is to generate findings that are valid and transferable [98, 168].

Validity refers to "the state or quality of being sound, just, and well-founded" [225]. The criteria for validity has been widely discussed for decades and while there is no universal agreement, a contemporary examination of approaches has defined the primary components as credibility, authenticity, criticality, and integrity [253]. I aim to achieve this by conducting data collection and analysis methods with rigor. I will conduct recursive and repetitive coding across the data corpus and iterate on codebooks to identify patterns and violations of such patterns.

Transferability refers to the extent to which findings can be utilized in other contexts and therefore a useful theory to apply in the world. Transferability is often used in contrast of generalizability, which is often the goal for quantitative research. Furthermore, within qualitative research, generalizing findings as universal is inappropriate because individuals and groups are unique and therefore urge examination from various perspectives [139]. Transferring findings is valuable because an understanding of any given setting may lead to insights that may explain relationships observed in similar or very different settings. For instance, Erving Goffman's impression management theory originally focused on within in-person interactions but has since been applied to digital contexts such as social media [81, 221]. Through dissemination of my main study findings at conferences, I hope to identify new domains where my findings are transferrable.

4 PRELIMINARY STUDY

This chapter presents my preliminary study in which I investigated how patient-generated data can be used in mental health clinical care. Specifically, I sought to address the following research questions

- **RQ1a** *What are mental health care providers' perceptions of self-tracked data?*
- **RQ1b** *What are the benefits and challenges of using self-tracked data in mental health clinical care settings?*

Below are the results of the study which was originally presented at CSCW 2019 titled *"Provider Perspectives on Integrating Sensor-Captured Patient-Generated Data in Mental Health Care."* I use the plural pronoun "we" to denote the contributions of the co-authors on this work.

4.1 Study Overview

Studies have shown that sensor-based patient-generated data (sPGD) have potential to improve health care by providing access to extensive information about patients' health and daily behaviors, thereby facilitating evidence-based treatment decisions. To date, there have been few studies of sPGD in mental health. It therefore remains unclear what opportunities providers see in using sPGD in a routine care setting and in relation to the physiological and behavioral manifestations of post-traumatic stress disorder (PTSD). Our investigation sought to expand our understanding of the potential role that sPGD might play in mental health care from the perspective of providers who are on the front lines of adopting these data into a new setting and navigating uncertainties around their interpretation and use.

4.2 Introduction

At an individual level, mental health disorders can cause social and cognitive limitations and lower quality of life, while at a societal level, they can reduce productivity and increase health care expenditures [40, 52, 53, 130]. Veterans experience elevated rates of mental health issues, especially PTSD. An estimated 23% of veterans returning from Operation Enduring Freedom/Operation Iraqi Freedom receive a diagnosis of combat-related PTSD [76]. Patients with PTSD tend to have high rates of comorbid mental and physical health conditions, such as depression, suicidality, and musculoskeletal and respiratory disorders [32, 72, 252]. Given the complexity of both PTSD and its frequently co-occurring disorders, individuals with PTSD often require ongoing management by health care providers.

One promising avenue for managing complex health conditions involves leveraging behavioral and physiological signals via wearable activity monitors like Fitbits. These devices have become increasingly ubiquitous since their debut a decade ago. A 2014 national survey found that one in ten American adults uses a wearable activity monitoring device on a daily basis [194], and that number is growing. Through sensors in the devices, activity monitors automatically track signals such as step count, heart rate, and sleep, on a continuous basis. Sensor-captured data from wearable devices is one form of patient-generated data (PGD) which has been increasingly considered for integration into different clinical practice settings to inform care. A growing number of studies have started to address the potential of sPGD as a tool to increase physical activity, for instance applying wearables as an intervention for overweight or sedentary patients [2, 17, 37, 84, 194, 228].

Applications of wearables to mental health have proceeded more slowly, reflecting the less straightforward mapping of available data streams to psychological states and symptoms. Yet, there are strong connections between physical and mental health [244], and signals detectable through sensors can also correlate with psychiatric symptoms. Early work exploring the application of wearables to mental health has suggested that wearables may support behavioral activation among patients with depression and motivate exercise among women with comorbid alcohol-dependence and depression, leading to reductions in anxiety and depressive symptoms [2]. In addition, studies of students and their college mental health providers suggest interest from both groups in applying sPGD to set personalized behavioral goals, assess therapeutic progress, and prompt reflection and discussion about the role of patients' behavior in mental health [125, 163]. While this represents a fledgling area of research, early findings suggest potential value of self-tracking for mental health.

Opportunities to integrate sPGD into mental health care also potentially align with recent calls for measurement-based care of mental health conditions [69, 91, 140]. In measurement-based care, treatment decisions are made through systematic tracking of patients' symptoms and responses to treatment. The guidelines that currently exist in relation to measurement-based care focus on self-reported assessments of symptom severity such as the Patient Health Questionnaire (PHQ-9) for depression [136], Generalized Anxiety Disorder Scale (GAD-7) for anxiety [224], PTSD Checklist for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (PCL-5) for PTSD [20], and Brief Alcohol Monitor (BAM) for substance abuse [28]. However, these self-report measures are subject to biases due to factors such as memory and social desirability. These biases do not affect sPGD. By using sPGD collaboratively with patients, providers may gain access to another stream of data to inform and adjust treatment.

Within CSCW there is growing interest in the collaboration required to leverage PGD generally in clinical environments [39, 164, 172, 205, 212]. We are still in the early adoption phase for sPGD, particularly in mental health care where providers are often slow to adopt health information technology [27], and where existing wearable devices have not been designed with mental health in mind. At present, we have little understanding of how mental health providers envision the role of sPGD in patient care and its potential impacts on their therapeutic role.

To address this gap, we conducted semi-structured interviews with 17 mental health care providers who treat veterans with PTSD and who represent a variety of specializations from individual psychotherapy to yoga instruction. These providers offered care in a three-week intensive PTSD treatment program in which veterans were given a complementary Fitbit. Although not required, members of the care team had the opportunity to engage with patients regarding the use of the Fitbit and its data. This study seeks to clarify how mental health care providers perceive applications of sPGD in this therapeutic context, as well as the barriers they perceive to realizing use of sPGD.

Our findings suggest that providers view patients' self-tracking with Fitbit as a means to promote empowerment and positive behavior change. They also envision using sPGD collaboratively, for instance in clinical conversations where data is situated in relation to both their clinical expertise and the patient's beliefs and experiences. However, given the climate of uncertainty around appropriate application of this novel type of health data, providers had enacted few of these envisioned collaborative uses. We discuss the challenge that mental health contexts may present to the "objective" nature of sPGD, the ways that perceived risks and uncertainty may drive treatment decisions in mental health care, and implications for designing systems to better leverage sPGD in mental health care.

4.3 Methods

4.3.1 Study Context: Post-Traumatic Stress Disorder

The lifetime prevalence of PTSD is estimated to be around 7% among American adults [129]. A diagnosis of PTSD requires both exposure to a trauma and a combination of symptoms clustering around intrusion, avoidance, changes in mood and cognition, and changes in arousal and reactivity [9]. Not everyone who experiences a traumatic event develops PTSD, and the development of PTSD is more common in women than men (20.4% to 8.1% respectively) and varies in response to the type of trauma experienced (1.8% for physical attack to 65.0% for rape) [128] Military personnel with exposure to combat stressors are at an elevated risk of developing PTSD, with an estimated 38.8% receiving a PTSD diagnosis. Especially when severe, PTSD can have negative consequences for life, work, and relationships [15, 128].

PTSD is often addressed in outpatient settings with evidence-based treatments including cognitive processing therapy (CPT). CPT focuses on discovering "stuck points," irrational beliefs about traumatic events. For example, a veteran suffering from combat-related trauma could believe they were the cause of others' untimely death through their commands. Dispelling stuck points involves Socratic dialogue and worksheets. In Socratic dialogue, the clinician asks the patient a series of questions to facilitate healthier reappraisal of traumatic events. In this process, the patient contributes their life experiences and interpretations while the clinician contributes their expertise on trauma and interventions. Clinicians may also assign worksheets for patients to complete outside of session. These may prompt patients to write and dissect narratives about the traumatic event [75, 197].

The uptake of evidence-based treatment for PTSD is poor due in part to patients' avoidance as well as difficulties accessing treatment [108, 123, 214, 249]. Some individuals with more severe or persistent symptoms also require a high level of care. One form of more intensive care is intensive outpatient treatment which condenses comprehensive therapeutic interventions into a short time period. Increasing the frequency of visits can lead to lower dropout rates and minimize between-session distractions that disrupt treatment [123, 124]. Intensive treatments are beginning to gain significant traction and yield positive results for a variety of mental health conditions [94, 211].

4.3.2 Study Setting: Intensive Treatment Program

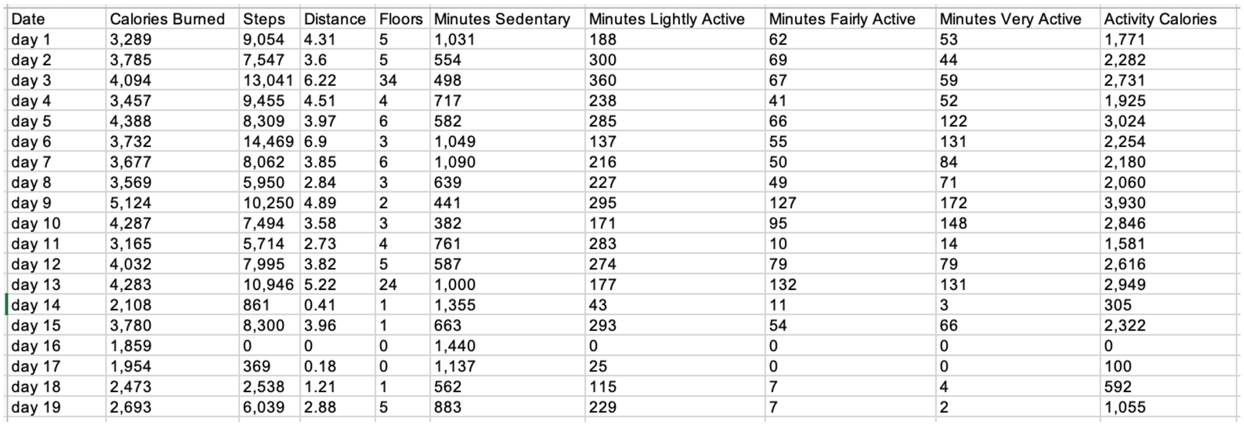
The Road Home Program: National Center of Excellence for Veterans and Their Families at Rush, funded in part by Wounded Warrior Project, created an intensive treatment program (ITP) in order to improve the accessibility, quality, and frequency of care for veterans living with the invisible wounds of war. This outpatient ITP lasts for three weeks and provides over 120 hours of comprehensive mental health care [citation anonymized for review]. Each cohort consists of 10-13 veterans diagnosed with PTSD. To measure progress over the three weeks, veterans complete four clinical survey assessments. These surveys track symptoms of PTSD

(PTSD Checklist for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, PCL-5) [20, 21], depression (Patient Health Questionnaire-9) [136], negative cognitions (Post-traumatic Cognitions Inventory) [67], and guilt (Trauma-Related Guilt Inventory) [137].

During the ITP, each veteran receives 13 sessions of group CPT, 14 sessions of individual CPT, 12 sessions of group mindfulness-based resilience training, and 13 sessions of group yoga. Each patient is assigned an individual provider, consistent across all individual sessions. Group providers are consistent across all sessions for each cohort. CPT providers include counselors, social workers, and psychologists. Mindfulness-based resiliency training and yoga groups are led by nurse practitioners, mindfulness instructors, and yoga teachers. In addition, the comprehensive team includes nurses, psychiatrists, neurologists, neuropsychologists, and psychology and social work trainees. In this study, we refer to this group as the "care team" and individuals as "care providers" as opposed to "clinicians" to encapsulate the diversity of expertise represented.

As part of wellness programming, Fitbits were incorporated into the fitness and nutrition courses to promote a healthy physical activity level and to facilitate self-awareness of the body's physiological responses to triggering events. On the day of the veteran's arrival, the staff provides the veteran with their own Fitbit and explains how to program it, use it, and interact with it, answering questions as the veteran sets up the device. Throughout the program, nutrition and fitness care team members incorporated the Fitbit into the curriculum through training and reviewing heart rate during sessions.

A prior study investigated how veterans in the ITP made use of Fitbit and its PGD [179], but it did not assess the care team's uses or perceptions of sPGD, as examined here. Prior to the application of Fitbit in the ITP, providers did not report extensive past engagement with PGD beyond soliciting patient completion of clinical symptom surveys; few reported engaging with behavioral or physiological data collected by patients outside the clinic. Care team members present at the time of the very first distribution of Fitbits received minimal (<1 hour) training on the data generated by the Fitbit. Team members that joined after this point were not trained, and most participants reported not being able to recall this training. Care team members did not have access to the Fitbit data unless the patient showed it to them through the Fitbit mobile application on their smartphone or through the Fitbit device itself. At the end of their three-week stay, Fitbit data for each consenting patient was downloaded in spreadsheet format once to store for potential future research (Figure 4). Care team members' largely undirected use of sPGD allows us to understand how routine mental health care processes may or may not be conducive to integration of these data.



Date	Calories Burned	Steps	Distance	Floors	Minutes Sedentary	Minutes Lightly Active	Minutes Fairly Active	Minutes Very Active	Activity Calories
day 1	3,289	9,054	4.31	5	1,031	188	62	53	1,771
day 2	3,785	7,547	3.6	5	554	300	69	44	2,282
day 3	4,094	13,041	6.22	34	498	360	67	59	2,731
day 4	3,457	9,455	4.51	4	717	238	41	52	1,925
day 5	4,388	8,309	3.97	6	582	285	66	122	3,024
day 6	3,732	14,469	6.9	3	1,049	137	55	131	2,254
day 7	3,677	8,062	3.85	6	1,090	216	50	84	2,180
day 8	3,569	5,950	2.84	3	639	227	49	71	2,060
day 9	5,124	10,250	4.89	2	441	295	127	172	3,930
day 10	4,287	7,494	3.58	3	382	171	95	148	2,846
day 11	3,165	5,714	2.73	4	761	283	10	14	1,581
day 12	4,032	7,995	3.82	5	587	274	79	79	2,616
day 13	4,283	10,946	5.22	24	1,000	177	132	131	2,949
day 14	2,108	861	0.41	1	1,355	43	11	3	305
day 15	3,780	8,300	3.96	1	663	293	54	66	2,322
day 16	1,859	0	0	0	1,440	0	0	0	0
day 17	1,954	369	0.18	0	1,137	25	0	0	100
day 18	2,473	2,538	1.21	1	562	115	7	4	592
day 19	2,693	6,039	2.88	5	883	229	7	2	1,055

Figure 4. De-identified Fitbit data from a patient that completed the program

4.3.3 Participants

Interviews were conducted in person except for one conducted by phone due to a scheduling conflict. All care team members in the ITP received an invitation to participate via e-mail from the research director and the clinical director. Participants were not compensated for their participation. Table 3 provides summary information for all participants.

Table 3. Participant Care Roles

Participant	Role
C1	Group therapist
C2	Individual therapist
C3	Non-therapist[b]
C4	Group therapist
C5	Group and Individual therapist
C6	Non-therapist[b]
C7	Individual therapist
C8	Individual therapist
C9	Individual therapist
C10	Non-therapist[b]
C11	Individual therapist
C12	Individual therapist
C13	Group and Individual therapist
C14	Individual therapist
C15	Non-therapist[a,b]
C16	Group therapist[a]
C17	Group therapist[a]

[a] Participant is also an organization administrator

[b] To anonymize participants who are the sole providers in their role, an administrator, nurse, nutritionist, and yoga instructor are listed as "non-therapist"

4.3.4 Procedure

In order to examine how health care providers in the ITP envisioned using Fitbit data, we conducted semi-structured interviews with 17 members of the program's care team. Three members of this group also served as administrators.

The interviews with the care team covered their: 1) role within the ITP, 2) experiences with PGD outside of the ITP, 3) perspectives on PGD, and 4) experiences with the Fitbit during the ITP. To gain a deeper understanding of possible uses of data outside of the current constraints of the Fitbit interface, care team members were also shown a list of broad categories of data Fitbit captures: heart rate, sleep, physical activity, consumption (food, drink), height, and weight. They were asked how they envisioned data of these types could be used within the context of the ITP.

Next, to constrain their envisioned uses to current technical capabilities and processes at the ITP, participants were shown a Microsoft Excel spreadsheet containing deidentified Fitbit data from a single patient who had previously completed the 3-week ITP (Figure 4). The patient data was chosen for its completeness over the three-weeks of the program in order to gather more reactions to the various types of available data. Care team members talked through their initial impressions of the data, the value they envisioned the data could bring, and how they imagined incorporating sPGD into their daily workflow. Participants were also prompted to envision having access to the data in real-time as the three-weeks elapsed, with completeness of the data depending on when the data is accessed. For example, if they were to check the Fitbit data on the third day of the program, the spreadsheet would only contain the first three rows of data. Finally, given that the care team can currently only access a patient's Fitbit data via the patient sharing data on their mobile app, we asked participants to explore a version of the Fitbit application shown on an iOS device. Care team members with administrative roles in the organization were asked additional questions about their observations of the care team's use of the Fitbits and their perspective on the role of the Fitbit within the goals of the ITP and within its parent program.

The Institutional Review Boards at Northwestern University and Rush University Medical Center approved this research.

4.3.5 Data Analysis

A third-party service was employed to transcribe audio recordings from all interviews. The resulting transcripts were analyzed by the first author through a process of open coding, which captured: current and proposed uses of data capturable by the Fitbit and barriers to such uses. Open codes were condensed into axial codes to create a code book, and all transcripts were coded using the resulting code book. The other authors peer-checked codes throughout. Taking a thematic analysis approach, the emerging patterns were sorted into themes that were derived from the data itself as opposed to existing theories [24]. This allowed us to identify patterns of use embedded in this specific context of care. Less prominent themes were disregarded, and the resultant themes were used to reanalyze the data set.

4.4 Findings

In this section, we present our findings on care providers' perceptions of the applications of Fitbit and its sensor-captured data to managing their patients' mental health. Providers were generally aware that their patients in the ITP were using Fitbits, and some reported specific interactions related to the Fitbit data. Providers also reflected broadly on the ways they envisioned sPGD being leveraged in the future, ranging from patient-driven Fitbit uses to active collaboration around the Fitbit and attempts to incorporate its data into routine delivery of mental health care. We organize our findings to reflect this distinction between uses that minimally or indirectly involve the provider and collaborative uses of sPGD.

4.4.1 Patient-Driven Uses of Fitbit and its Data

While providers lacked extensive knowledge of how their patients were using the Fitbit during or after the ITP, they envisioned a number of ways in which patients could monitor and reflect on patterns in their Fitbit-captured data, leading to improved understanding and management of their own health. Several had examples of discussing these data in care when initiated by the patient. However, providers also noted potential challenges of interpretation that could arise in patient-driven use.

4.4.1.1 Patient-Driven Uses of Fitbit and its Data

Providers largely agreed that patients' uses of Fitbit might involve monitoring of and reflecting on patterns in their data (e.g., step count, heart rate, sleep logs) as objective records of their activity. This monitoring was generally viewed as a source of patient empowerment, and some had seen evidence of such benefit in practice. For example, C16 described the sense of validation a patient had experienced by reviewing patterns in her data. This patient had believed her feelings of exhaustion were inexplicable, given their persistence despite efforts to get reasonable hours of sleep, but she discovered that the Fitbit tracked not just hours in bed but also the quality of sleep. This metric yielded a new insight, providing a better account of her troubling symptoms. C16 described:

> *"It's almost like getting an outside opinion... from the Fitbit that helps them see, that helps them better understand their symptoms… That's really validating." (C16)*

Providers also speculated that patients might be motivated through accessing a record of their accomplishments, such as meeting step count goals. As C16 continued:

> *"I think that the benefit for the step challenge, for example, is obviously getting physical exercise, likely more than they have in the past because they are kind of prompted to do more. It also helps them feel a sense of accomplishment." (C16)*

This validation might be particularly valuable to patients with PTSD whose sense of self-worth is often impoverished:

> *"A lot of the thoughts you have are that you are incapable, inadequate, cannot accomplish things. So, that [data] kind of directly speaks against that, right? 'I am able to accomplish something, like reaching 10,000 steps a day.'" (C16)*

C4 described seeing these benefits in practice, with patients reporting that they had used sPGD to counter some components of PTSD, such as negative beliefs. In particular, some patients who had been isolated and inactive because of PTSD expressed their delight when increases in activity became apparent through reviewing sPGD. C4 reports that, in response, *"I just try to provide affirmations for the change in their behavior."*

In addition, some providers hypothesized that sPGD might even surface benefits obtained through therapy, such as a change in symptoms after completing a written narrative of their trauma. For example, C4 described that accessing Fitbit's graphs of symptoms could hypothetically provide a patient with:

> *"…a visual of progress, or how things are changing over time for them. It can connect. For example, if somebody writes their trauma account… and then, they notice a change in how they slept." (C4)*

In these examples, the data from Fitbit have value in allowing patients to self-monitor. They also have potential to provide an outside perspective for patients, which might yield new insight or validate a patient's effort and therapeutic progress. Indeed, some of these benefits had already been reported by patients in clinical conversations.

4.4.1.2 Reinforcing Negative Views and Agitating Symptoms

Providers also recognized situations where patients' use of the Fitbit could lead to negative consequences. For instance, PTSD can be characterized by the tendency to catastrophize or become hyper-focused on negative occurrences. Several providers reported concerns that, upon reflecting on their data, patients could become discouraged by apparent lack of progress, or might focus on negative trends:

> *"Sometimes it can have an alternative effect like, 'Yeah, look how shitty I'm sleeping, no wonder I feel so bad' and then it can create a focus on negative things as opposed to not being aware that you're sleeping so bad. It creates something else like, 'Oh, look at me, I'm not even sleeping well. It's just one more thing like life sucks' kind of a thing. So, I wouldn't say that that's typical but that's also a way that it can actually have a negative impact… if somebody is overly negatively bleak about their life and everything around them then it could just be one more thing to reinforce how bad they have it." (C14)*

Providers also reported that overuse of Fitbit might provoke anxiety in patients. This had already occurred with at least one of C9's patients, such that C9 had to shift focus away from the data to a more productive task. He described:

> *"People get a little obsessed about sort of tracking symptoms. I had a client that I can remember that wore his Fitbit and every time he would get anxious he'd be like, 'My heart rate's at this, my heart rate's at this,' and that would just get him really amped up and he would obsess about this and it was like this instant biofeedback which just really elevated his anxiety because he was just so fixated on his heart rate and so we actually had to say, 'Let's dial back on you monitoring every few minutes. It's useful information to have but let's refocus on what we're doing here.'" (C9)*

In C9's view, not only did the patient's use of sPGD reinforce anxiety, but conversations about the sPGD led to loss of time that might have been spent on evidence-based practices like CPT. These examples suggest that, from the provider's perspective, the same data from the Fitbit might either support empowerment or induce fixation and anxiety depending on context, suggesting an ambiguous relationship of Fitbit data to patient wellbeing. C11 provided an example of how productive or unproductive application of Fitbit data may reflect a particular patient's approach to coping:

> *"I think physical activity can go both ways. It's a coping skill for a lot of people, but it can also be a form of avoidance.... So, 'Oh, you were distressed today and then you went and ran so hard that you made yourself puke? … Is this healthy exercise or is this you not wanting to feel your feelings?' I know that that's come up with a couple of my veterans as far as limiting their physical activity to, 'Okay you're going to go work out for an hour, but you're not gonna push yourself to a point of pain because that's avoidance.' So, depending on the person, sometimes that's real useful information. Other veterans will be like, 'Oh, I've been walking to and from group today. It feels really good. I've been getting my steps in and it's been really helpful. I've noticed a difference.' That's helpful." (C11)*

This situation highlights the importance of re-contextualizing physiological data within a clinical process where a provider may intervene to suggest an alternate way to frame or interact with the data. The provider may choose to de-emphasize negative patterns or to steer a patient away from data fixation or avoidant behaviors that have potential to be exacerbated by self-tracking, such as over-exercise. In the next section, we describe how providers recognized potential benefits of engaging with sPGD as a routine aspect of treatment.

4.4.2 Integrating Fitbit Data into Treatment Protocols

While providers reported that patients occasionally discussed their Fitbit data in clinic visits, providers largely had not pursued integration of the Fitbit and its data into their routine mental health care. However, providers envisioned that the Fitbit and its data could support their work by capturing additional information outside of the treatment setting, even providing a new means of systematically testing effectiveness of treatment options in order to improve care. When assessing a patient's condition, therapists currently rely on three streams of data: (1) observation of the patient's behaviors and interactions, (2) the patient's self-report, and (3) the patient's scores on various survey measures. Analyzed together, providers can identify areas of consensus or discrepancy between the different data streams, sometimes highlighting a need for discussion and clarification. For example, if the patient's PHQ-9 scores decrease (i.e. patient's depression is decreasing) yet the patient appears to be withdrawn during group therapy sessions, there is a conflict between the two points of data that needs to be reconciled. Some providers imagined that the Fitbit data could serve as a fourth stream of data that could offer triangulation with other available data streams and also provide ongoing insight into the patient’s status and behaviors outside of the clinic.

4.4.2.1 Understanding the Relationship Between sPGD and Standard Measures

In order to make this new stream of data useful, providers stated that sPGD should ultimately connect back to survey measurements. C17 speculated that it might be possible to identify a relationship between PTSD symptoms as measured by surveys and behaviors tracked by the Fitbit, potentially even allowing a view into therapeutic progress across the ITP:

> *“What I think we could do more of with the Fitbit is seeing [patterns of sleep]…what was their baseline for the first week and then going into the second week what’s the average there and then the third week and seeing if that mimics the scores that we see on their self-report measures. Usually people will be very distressed in the first week, experience significant hypervigilance, an increase sometimes in frequency and intensity of their symptoms and then it will hit a peak and then will decrease into week three so I would be curious as to whether the Fitbit would show that.” (C17)*

Having observed a common trajectory of symptom scores among patients undergoing three weeks of treatment, C17 was curious about whether the Fitbit could also surface these patterns.

4.4.2.2 Integrating sPGD with Therapy

Beyond triangulating patterns observed in other data streams, care team members envisioned that the Fitbit could be used to aid in treatment, such as through facilitating “reality testing” to help patients overcome false or dysfunctional beliefs. C1 envisioned a situation in which the step count feature could be applied to identify negative cognitions and dispel them during individual therapy sessions:

> *"If somebody's steps are really low for the entire ITP, it might lead to a discussion about 'What's that about?' and then it might elicit a belief or cognition of, 'The world is not safe' or, 'If I leave my room something terrible might happen.' That could result in a homework assignment being given about, 'Why don't you go out with the group tonight and just see how it goes? Experiment.'" (C1)*

Thus, accessing the Fitbit and its data could prompt dialogue around a stuck point centering on the consequences of being out in the world. The Fitbit could then be used to quantify or track progress toward an increased activity goal as the participant challenges the stuck point.

The Fitbit data was also viewed to have the potential to guide more productive discussions during treatment. For instance, C2 described how discovering correlations might open up a conversation about causes of stress:

> *"It could be beneficial [to] have that information and say, 'Oh, wow, as your scores have gone down, your sleep is gotten better' [or] 'As your scores have gone down, your heart rate seems a little more even' [or] 'Your heart rate is always up during the group, what's going on with that?'" (C2)*

While these uses were hypothetical, providers were optimistic that the signals tracked by Fitbit might offer new opportunities to engage in therapeutic dialogue. In particular, observing correlations between symptoms and behaviors was envisioned as a potential way to spark conversations about progress and changes throughout therapy.

4.4.3 Challenges to the Use of sPGD in Treatment

Providers also anticipated potential negative consequences of collaborative activities involving sPGD, such as diverting time from evidence-based practices. They highlighted that displacement of standard care could negatively affect the patient's treatment and cost the organization valuable time and resources.

4.4.3.1 Connecting to Clinical Practice.

Especially given the present context wherein patients decide whether and how to use the Fitbit and which data to share with providers, providers questioned whether they would have opportunities to connect data with their domain expertise: evidence-based care, which in this program was CPT. Providers wanted not only evidence that use of the Fitbit would significantly help patients, as measured by self-report surveys, but also training to help them connect Fitbit data to their established clinical practices. C8 outlined these concerns about not adhering strictly to CPT:

> *"It's noticeable if a clinician is not sticking to protocol because, a) the patient's scores aren't changing and they're not as engaged in the program and the client lags behind and, b) If I were like, 'We went into the app and talked about sleep data' we would be asked 'Why aren't you sticking to the protocol?' not in a mean way or not in a punishing way but in a way that's just like everyone needs to be a united front and consistent." (C8)*

Given that patients' health is at stake, providers in the ITP highly value providing care that is effective and cohesive. It was not yet clear that the Fitbit and its sPGD would help in this effort.

Questions about the validity of the sensors were also cause for concern. C1 hypothesized that inaccurate readings from the Fitbit could cause unnecessary alarm and waste time in "non-productive conversations," derailing clinical practices and undermining protocol. For example, this was imagined to occur if the Fitbit indicated poor sleep in a patient who in reality slept well:

> *"The Fitbit sleep data is not as solid as other sources of data so I could see it backfiring somewhat…that if a person is actually sleeping fine their Fitbit could say that they're not or vice versa it might just confuse things." (C1)*

Thus, measurement errors might be counterproductive to treatment, even potentially leading to an exacerbation of patients' issues.

4.4.3.2 Time and Resources

Another concern regarding integrating Fitbit into care was the time and the resources needed to incorporate new streams of data. As C7 described:

> *"There's always so much to do.… You're already kind of preparing for sessions, and at a certain point it's like, 'How much am I treating or assessing what I'm seeing on a screen or on paper compared to just talking to someone and figuring stuff out together?'" (C7)*

In this quote, C7 highlights the potential that time needed to understand sPGD might negatively impact a meaningful interaction with the patient.

Furthermore, if Fitbit data were to be integrated into the clinical data repository (as opposed to the patient approaching the provider to share data), participants noted that additional individual preparation work might be required in order for patient-provider collaboration to occur. To prevent wasting time in sessions, the care team members envisioned reviewing the data before the patient comes in in order to decide whether there is something notable to discuss. However, as C5 notes, therapists may be unclear on how to recognize what is worth discussing and, furthermore, how to act on it.

> *"You're going to need [to] train them on how to balance all of these different pieces of data they have access to and how to prioritize the data. I think it would be especially important for new therapists coming on. It would probably be pretty overwhelming for some to have access to that much data and I think we would need to do like a standard operating procedure of how to [deal with] the information." (C5)*

Providers felt they had relatively little flexibility in adopting new practices given the organization's emphasis on protocols to ensure standardization and cohesion. Administration of specific symptoms survey measures was directly tied to grants and funding sources for the organization. As one provider and program leader (C16) stated:

> *"If our main funders want to see you doing a program for veterans with PTSD, what we want to see is reduction in PTSD symptoms and obviously that's what we want to focus on. If we had another funder that said, 'Hey, we'll give you this much money to make sure that all the veterans with PTSD see an improvement in their sleep or their exercise' or any of that, then our program would be structured very differently and maybe would incorporate more of these types of information.… I don't think the goals will change simply because they are pretty set on what they're looking for, but if they were to change, I think that would directly impact what we do in here because we specifically chose CPT because it's an evidence-based treatment for PTSD because that's our goal." (C16)*

At present, there are no specific evidence-based practices for the treatment of PTSD that involve data the Fitbit can track. Consequently, providers noted a lack of organizational incentive to use the data, especially given its unknown value in treatment and the time and resources that would be required for training and integration.

4.5 Discussion

Our study examined providers' perceptions of the application of sPGD to delivering mental health care. In current practice, Fitbit and sPGD use were self-directed by patients, and providers recognized a number of ways that sPGD could help patients to generate new insights about symptoms, provide external validation, and reinforce therapeutic gains. At the same time, providers also expressed concern that patients might interpret data in unproductive ways or fixate on negative patterns. Clinical interactions were viewed as potentially fruitful settings within which to contextualize sPGD in relation to the patient's experience and goals, while also drawing on the provider's therapeutic expertise to interpret data and extract insights relevant to treatment. Nonetheless, given the limited validation of sPGD as clinically relevant measures in mental health and the lack of protocols to accommodate collaboration around these data into clinical workflows, providers had done little to apply Fitbit as a mental health tool beyond offering consultation when initiated by patients. These findings echo our prior work with veterans completing the ITP, which found that the Fitbit was primarily used for self-monitoring and supporting social interactions with peers, rather than collaborating with the care team [179].

The gap we describe between providers' envisioned and actual collaborative uses of Fitbit and its data reflects a number of challenges these providers face in adopting novel technologies within a mental health care environment. In our discussion, we describe that despite the promise of sPGD as a means of bringing "objectivity" into mental health care, applying sensor-captured data involves a complex and uncertain interpretive process that, from the provider's perspective, is imbued with potential risk. In the sections that follow, we situate our findings in relation to prior work highlighting the subjectivity inherent in extracting meaning from technology-based measurements and we discuss organizational factors relevant to negotiating risks emerging from this interpretive process. Finally, given the unique constraints of sensor-based data within this mental health space, we identify design opportunities for patient-provider collaboration around sPGD.

4.5.1 Shifting Fitbit Data from “Objective” to “Situated Objectivity”

Our data highlight a tension between the perceived authority of sPGD as a source of “objective” measurement and the subjectivity inherent in extracting insights from these data. Below we describe how a process to appropriately situate data is necessary in order to establish meaning so that sPGD can be effectively used in mental health services.

4.5.1.1 Situating Objective Data in Mental Health

Pantzar and Ruckenstein have observed that there are prevailing perceptions of tracking devices as a source of “mechanical objectivity” [47] where, due to their automated and standardized collection of data, their measurements are perceived as essentially objective [186]. Yet, they argue that the meaning of sensor-captured data is in fact deeply tied to the particular contexts in which data are collected, and data must therefore be appropriately situated to offer reliable insight. For example, sensor-captured data might indicate a user’s elevated heart rate over their baseline in a certain timeframe. Yet, there could be multiple possible reasons for this pattern; elevated heart rate might reflect stress and motivation as an individual strives to achieve a meaningful goal, but it could also reveal anxiety that impairs that individual’s ability to function. In this example, contextualizing sensor-based data within an individual’s larger experience is central to interpretation. Pantzar and Ruckenstein therefore propose the concept of “situated objectivity” to describe the epistemological status of sensor-tracked data, combining the authority of mechanical objectivity with the important role of context in knowledge formation.

Our findings suggest that providers view clinical interactions as a promising site for developing situated objectivity, with both patients and providers contributing valuable context. Specifically, clinical conversations may facilitate elaboration by the patient on her subjective experience during data collection, and on her initial assessment of the meaning of these experiences. In turn, the provider might suggest an alternative framework through which to view these experiences, drawing on a therapeutic perspective. For instance, Fitbit data might show the patient's elevated heart rate whenever she is in a particular location. The patient might identify that this physiological reaction reflects an experience of anxiety. Without guidance from the provider, the patient might take this as a sign to avoid that location to prevent this response. However, from a therapeutic perspective, such avoidance might be unproductive, failing to address the cause of the anxiety. The provider might recommend against avoidance, instead suggesting continued visits to the site while employing coping methods learned in therapy, allowing the patient to ultimately overcome the anxious response. In this example, despite the application of non-traditional data to spark the conversation, the therapist operates within his or her therapeutic expertise by proposing a new possible lens through which the patient might view her experience in order to facilitate the development of mastery and resilience [71].

4.5.1.2 Overcoming Interpretive Challenges

Our findings also suggest that providers see a potential role they could play in helping patients overcome specific interpretive biases that might manifest in relation to sPGD. One challenge among patients with PTSD is a tendency to focus on negative aspects of experience at the expense of positive ones [127]. In this context, the perceived "objectivity" of sensor-based data can play a dual role. On the one hand, providers hoped that sPGD could reduce the bias involved when patients monitor their symptoms and health behaviors in daily life, speculating that these data could offer an "outside" perspective, potentially more tethered to reality, and more valid and reliable than patients' perceptions. This authoritative perspective from sPGD was viewed to have beneficial effects when it allowed patients to recognize their efforts and accomplishments (e.g. achieving step count goals). However, by the same token, sPGD could authoritatively reveal negative patterns of behavior or lack of progress in ways that might be demotivating. Thus, perceptions of sPGD as an unbiased signal could enhance or diminish motivation, largely reflecting the particular patterns the patient extracts from the data. Given the numerous potential patterns that patients could focus on within the Fitbit data, the input of the provider may become especially valuable to counteract negative bias, including by highlighting the positive trends and accomplishments, or recommending against over-tracking in those who become fixated.

Our findings present how mental health providers may have an important role to play in situating sPGD, including reframing and refocusing data interpretation. These findings echo those of West et al. [251] who found that mental health providers perceived their patients to need help overcoming the substantial ambiguity of self-tracked mood ratings. However, it is worth noting that providers' perspectives offer value in contexts beyond mental health, with work in physical health also suggesting the value of collaboration to overcome data ambiguity [164, 165, 172]. For instance, similar to our findings, Mentis et al. [165] describe "cointerpretation" that occurs as a negotiation between perspectives of patients and providers in clinic visits. Yet, as the next section describes, despite value of multiple perspectives for interpreting sPGD, providers in this setting faced barriers to taking on this role in routine practice.

4.5.2 Uncertainty and Risk of Non-Traditional Mental Health Data

In this section, we describe the barriers to integrating sPGD into mental health care, focusing on the ways providers are inhibited by uncertainty in deciding which measures to examine, lack of clarity on what sPGD-driven actions to take, and concerns that emphasis on sPGD might displace validated measures and better-established therapeutic activities.

4.5.2.1 PGD as an Unvalidated Measure

In a commercial context, wearables provide data that may sometimes lack precision and accuracy, but that still generally fall within an acceptable "uncertainty tolerance" for consumers [133]. However, providers using such devices, including those in mental health settings, may have substantially lower tolerance for risk. Indeed, managing risks and benefits for their patients is a large part of providers' jobs, and prior work suggests that providers are often slow at adopting new technology because of concerns about liability and risk [93, 135]. To manage risk, providers typically rely on evidence-based practices, which include using validated measures and assessments. Commercial devices that passively capture PGD, such as Fitbits, have yet to be empirically validated for many of their potential uses in health care, particularly in mental health care. In physical health, many self-tracking technologies at least mirror "gold standard" analog tracking practices, such as blood glucose monitoring, calorie counting, and heart rate variability. In contrast, in mental health, sensors currently map onto symptoms like stress and depression in ambiguous ways, with "gold standard" measures generally coming from validated self-report instruments [69]. While a body of evidence is emerging to support correlations between sensed data and self-reported symptoms, such models generally combine an array of signals [173]. This work has also relied heavily on "black box" algorithms that humans cannot easily understand [18]. At present, there is limited evidence that care team members can manually extract clinically meaningful information from self-tracked data from commercial wearables.

In medicine, risks of a new measure or approach can be justified at times based on potential clinical benefits. Yet, benefits were viewed as uncertain in this study. While providers suggested possible therapeutic uses for sPGD, they also emphasized that benefits would be inconsistent across patients. With clinical relevance of sPGD to mental health still unclear, especially in contrast with evidence-based practices, it is unsurprising that providers were hesitant to engage with sPGD.

4.5.2.2 Organizational Fit & PGD

While the providers described in this chapter are on the frontlines of this novel application of PGD to mental health, a number of organizational factors make them unlikely evangelists for new technologies.

First, these providers operated in an organizational context in which their workflows and incentives were closely tied to self-report measures of mental health symptoms that were not associated with Fitbit. In this context, if a provider dedicates time during a clinical interaction to interpreting sPGD, they risk disrupting the organization's established practices. Such disruption could even come at the cost of patient health, since focusing on sPGD takes time and attention that might otherwise be applied to evidence-based treatment. Second, if a patient's health does not improve per symptom questionnaires, the provider's care process could even come under scrutiny, a possibility mentioned by at least one provider in this study. In general, providers wondered whether there was organizational buy-in for their role in supporting patient use of the Fitbit such that they would be appropriately trained and accommodated in taking on the substantial work and risk involved. Finally, when we consider health information technology, it is often important to evaluate innovations post-implementation to understand how workflows and other organizational features affect adoption [44]. However, sPGD creates an additional challenge because of the potential use of the collected data in the care process itself. With the growing emphasis on evidence-based practice in health care [68], administrators and policymakers will need to ensure that there is a strong evidence base supporting safety and effectiveness prior to even implementing sPGD technologies. Otherwise, providers may be reluctant to use these tools. Other studies have also highlighted unmalleable policies, introduction of new risks, and lack of incentives for adopting new work practices as major barriers to implementation [176, 183, 190, 245].

While these challenges to adoption are significant, CSCW and Implementation Science researchers highlight how we could begin to navigate organizational barriers [41, 66, 80]. For example, the CSCW community has highlighted the importance of understanding organization members' interpretive frames for making sense of emerging technology [183] and the

Implementation Science community has emphasized how willingness to adopt an innovation can reflect positive characteristics such as trialability, which may involve providing low-risk situations within which to experiment and become comfortable with new technologies [48, 63, 86]. This may also include the provision of training, a social climate where key stakeholders support social influence [46, 131], and where organization members can learn from one anothers' experience [35]. The success of sPGD in healthcare organizations will depend on ensuring the organizational readiness to adopt these tools.

4.5.3 Design Considerations for sPGD in Mental Health

In this section, we describe challenges designers face in supporting use of sPGD in mental health, including in facilitating data interpretation by both providers and patients. We also discuss considerations that may emerge as mental health states become more reliably linked to sensor-based data.

4.5.3.1 Designing for Multiple Perspectives

Our findings suggest that providers and patients have different interpretive approaches and priorities for sPGD, with providers emphasizing opportunities to inform therapeutic activities, and patients looking to validate their experiences, enhance self-understanding, or motivate themselves outside the clinic. These differing patterns of interacting with the data suggest a need for different types of interfaces for different stakeholders through which to access, review, and interpret sPGD.

In considering possibilities of using sPGD in their work, mental health providers emphasized their need to constrain the application of sPGD to data considered more relevant and familiar, and for which the organization holds them accountable. Similarly, other studies have found that clinicians want patients to track data considered “clinically relevant” [251, 259]. Yet, in the present context, Fitbit use by patients is voluntary and self-directed, and providers have little control over what data patients track or share. To support extracting relevant information, providers will likely require more comprehensive, reliable, and timely access to data. While fundamental, this step is not simple. Prior work from a physical health context suggests barriers to data sharing including privacy concerns as well as incompatible formatting and challenges incorporating data within electronic health records [39, 251, 259]. Furthermore, having accessed data, providers likely need support to extract meaningful signals. This could be facilitated by automatic pattern recognition, for instance, generating alerts based on changes in physical activity as captured by wearable sensors [31], and changes in health status as captured through passive sensing in a housing facility [222]. Pattern recognition and interpretation can also be facilitated through automated data annotations, such as natural language captions for correlations between different self-tracked measures (e.g., sleep and stress) [18]. In this work, the presence of captions increased users’ understanding of behavioral patterns and of these patterns’ potential consequences. Visualizations may also help to make correlations between various signals salient (e.g., through bubble and bar charts) [212]. Finally, as Raj et al. [193] note, providers’ appropriate response to changes in patient status may require access to contextual information. For instance, providers may benefit from labels that specify whether data were collected when patients were at home, at work, or in the clinic.

Considering the needs of patients suggests additional design considerations. In particular, providers suggested negative consequences when patients over rely on quantified metrics in monitoring their health status, expressing concern that patients might use data in ways that manifest PTSD and its related symptoms by becoming fixated on the data or using it to harshly evaluate themselves. Such considerations could be addressed, in part, through the ways data is visualized. Researchers and designers in the HCI space have previously explored departures from numeric presentation of tracked data towards visual representations that allow for subjective interpretation and that prompt reflection and mindfulness. This has included imprinting physiological data on artifacts, such as when Howell et al. [105] created a shirt with a colorful display corresponding to emotion as detected by changes in the wearer's skin conductance, and when Thieme et al. [232] created spheres that change color based on the holder's heart rate. As for non-digital data representation, Snyder et al. [223] worked with individuals with bipolar disorder to create speculative visual representations of their lived experience. Although these examples might not translate into a clinical environment, they highlight a shift from quantified PGD to instead emphasize multiple possible meanings of data, potentially reducing fixation with quantified self-monitoring, as observed in this and other studies [125, 150, 212]. These potential benefits should be weighed against the usefulness of numeric representations in identifying trends and anomalies.

Design may also facilitate collaboration and integration of multiple perspectives. For example, Chung et al. [39] show that patients and providers rely on data artifacts to facilitate sharing and contextualizing self-tracking data. Through these data artifacts, patients and providers can prioritize topics of discussion to cover during clinic visits [212]. Each party also can draw attention to patterns that they find important, offering tentative accounts of the data's meaning, perhaps through highlights or annotations [164]. Options to annotate data may be especially desirable in mental health, where patterns in sPGD cannot be interpreted without understanding the wearer's subjective mental state. Yet, given the numerous possible correlations to draw from the data, it may be helpful to constrain the annotation process. For instance, automated pattern recognition could be applied to identify a subset of patterns and anomalies from sPGD which the patient might then annotate, situating them in subjective experiences such as motivational states and occurrence of symptoms.

4.5.3.2 Leveraging Sensor Data in the Future

This study investigated the ad-hoc use of a commercial wearable in a routine care setting, expanding a body of research on the clinical use of passively generated data in mental health. While our findings suggest a high level of ambiguity perceived in the signals from a commercial wearable in this care setting, this in part reflects that Fitbit and similar wearables focus on physical rather than mental health sensing. This will likely change in the future as research and development advance to better detect and predict mental health states through passively collected data, sometimes in combination with self-reported data [173, 239]. Much of the work in this area has leveraged smartphones rather than wearable devices, including smartphone-based sensing of geolocation, vocal quality, accelerometry, communication logs, and social interactions as based on co-presence of other devices and sampling of ambient noise [14, 62, 115, 173]. Other work has explored using physiological measurements not routinely available in commercial wearables, such as electrodermal activity [78]. In addition, digital trace data, such as social media posting content has been applied to predict mental health states [204]. This work is still at an early stage and largely outside the awareness of the providers interviewed in this study. However, integration of these emerging methods into treatment and management of mental health must ultimately be informed by collaboration with the stakeholder communities these tools will impact, including patients, who will be asked to allow these devices to collect and transmit their personal data, and clinicians, who will be required to use this data to inform and adjust their clinical care. In addition to concerns related to impacts on clinical workflows and evidence of clinical efficacy, recent work on passive tracking has also identified ethical concerns with privacy and monitoring of populations [199].

4.5.4 Next Study

This preliminary study, which focused on mental health care providers, highlighted new opportunities for research and design of self-tracked data. In particular, shifting the notion of self-tracked health data as information that is *mechanically objective* to information that is *situationally objective* introduces new aspects of self-tracking data interpretation and design to explore in future research.

Regarding data interpretation, our study highlighted the importance of understanding context in knowledge formation. The context we studied, an intensive treatment program, was not the patient's routine environment, either physically or socially. As one care provider (C1) in stated:

> "*Part of me wonders…They've traveled here. This is totally not their home environment, you know, a lot of them are from more rural areas and now they're in the city. There's just a lot of reasons why their activity level and sleep and heart rate would be different*

> *than it is typically so I'd be interested almost to getting [ITP] baseline data while they're here and then when they go home for like three or six months. Like…before coming and then after they finished the [ITP] and go back home – are those different?"*

The intensive treatment setting can be conceived of as a disruption, a non-routine circumstance of the veterans' lives. In the main study, I sought to gain a better understanding of the perspective of self-trackers themselves and how they perceived their data collected during a period that exists outside of their routine lives. Through this research, I answered the following research questions:

- *(RQ2a) How do self-trackers define which data captured non-routine circumstances are useful to reflect upon?*
- *(RQ2b) How do self-trackers envision reflecting on data after the non-routine period has passed?*

Once I gained an understanding of how to support individuals reflecting on data from non-routine circumstances, I also asked:

- *(RQ3a) How should we think about designing data visualization-modification features that support the reflection of data after the non-routine period has passed?*
- *(RQ3b) What are the benefits and challenges of modifying data visualizations that depict a non-routine period?*

In the next chapter, I present an empirical study that investigated the self-tracker perspective on their data captured during such a circumstance and how to design for them.

5 MAIN STUDY

5.1 Study Overview

In the preliminary study, I examined the use of wearables by mental health care providers in the treatment of their patients during a non-routine circumstance: an intensive treatment program. This study highlighted how the clinical perception of data is influenced by the understanding that the data were captured during a period outside of the patient's routine life. The care providers were uncertain whether data captured by the wearables were a reflection of therapeutic progress or of the patient being in an unfamiliar environment. They perceived an ambiguity as to whether the behavioral data captured during the treatment program will have any indication on the patients' behavior once they return to their routine environment. These results suggest that breaks in routine influence how data captured during these periods are interpreted by individuals reflecting on data.

In this chapter, I describe a two-phase study to further explore this concept. In this study, I interviewed women on their perceptions of data they collected during their recent pregnancy, another instance of a non-routine circumstance (see Table 4).

- **Phase 1**: I conducted a remote qualitative study that employed semi-structured interviews which allowed participants to express their own conceptualizations of the utility of data captured a non-routine circumstance and a textual-elicitation activity to further elicit their values [16]. I examined how women perceived of self-tracked data captured during their pregnancy, its utility, and the temporality of its relevance. I answered:

  - (RQ2a) How do self-trackers define which data captured during non-routine circumstances are useful to reflect upon?
  - (RQ2b) How do self-trackers envision reflecting on data after the non-routine period has passed?

- **Phase 2**: Based on interviews from Phase 1, I identified three data visualization-modification features that participants may want to employ when reviewing data from their pregnancy. These were the ability to: (1) *archive* data, (2) *highlight* segments of data, and (3) *mark* up graphs. During each interview, I presented three graphs of simulated health data: (a) weight, (b) hours of sleep, (c) daily stress ratings, and asked them to pretend this data was captured during their pregnancy. I probed each participant on how they might want to modify each graph (*archive, highlight, mark,* or *other*) and how these modifications would support their self-reflection. This empirical investigation on how these three features may support self-reflection answers the questions:
  - (RQ3a) How should we think about designing data visualization-modification features that support the reflection of data after the non-routine period has passed?
  - (RQ3b) What are the benefits and challenges of modifying data visualizations that depict a non-routine period?

Table 4. Summary of main study phases, methods, and research questions

Phase 1 Need-finding _Research Methods_ *Semi-structured interviews with textual-elicitation activity*	RQ2a How do self-trackers define which data captured during non-routine circumstances are useful to reflect upon? RQ2b How do self-trackers envision reflecting on data after the non-routine period has passed?
Phase 2 Design evaluation _Research Methods_ *Semi-structured interviews with visual-elicitation activity*	RQ3a How should we think about designing data visualization-modification features that support the reflection of data after the non-routine period has passed? RQ3b What are the benefits and challenges of modifying data visualizations that depict a non-routine period?

5.2 Introduction

Efforts to facilitate the integration of measurement-based care into routine care processes often result in the development of technology that passively capture signals related to health, but neglect broader relevant life contexts that may require manual tracking of signals [140]. For instance, a team of neuroscientists, engineers, and researchers at Project Amber sought to find biomarkers for depression through electroencephalography data: electrical activity in the brain. Ultimately, they concluded that while their metrics have potential to aid in clinical decision-making, the value of their tool lies within the confluence of biometrics and the patient's subjective experiences of daily living – the tool on its own, was ineffective[7]. This is not surprising, as context has always been vital to the interpretation of health data. When it comes to diabetes management, it is critical to know how often a patient is physically active and their eating habits. For depression management, identifying patterns of movement and social

---

[7] https://blog.x.company/sharing-project-amber-with-the-mental-health-community-7b6d8814a862

engagement with respect to time of day or duration can inform timely interventions. Inversely, data taken out of context can be uninterpretable or worse, misinterpreted.

The increasing rate at which individuals are generating data on themselves using commercially available tracking devices means more data is being collected throughout more varied life contexts than ever before. Individuals can now possess data from decades ago and soon, their entire lifetime as parents provide their children their own wearables [257]. It is foreseeable that such digital artifacts may eventually be preserved for multiple lifetimes or even generations [150, 166]. With the proliferation of Internet of Things devices, interconnected data can be used to identify patterns beyond the individual level to a group or societal scale. Low-burden, passive context detection is an increasingly pertinent area of research as technologists begin to expand the uses for these data [6]. Incidentally, these data will inevitably be captured during periods of disruption or transitions between routines.

In this two-phase study, I explored women's perceptions of self-tracked data captured during non-routine periods such as pregnancy. In the first phase, I conducted semi-structured interviews and used text-based elicitation to understand how women envisioned using (or not using) data that was collected during pregnancy and other periods that they do not consider their current routine. In the second phase, I used visual representations of simulated pregnancy data to evaluate various data visualization-modification features and as a probe to understand how women define boundaries within their data. The participants were comprised of 21 women (Phase 1: n=8, Phase 2: n=13) who had recently given birth and had tracked their stress during their pregnancy. I found that depending on the type of self-tracked data collected, participants implicitly perceived boundaries in segments of data, what I refer to as bounded situational contexts (BSCs). These boundaries were defined based on an individuals' internal contexts (e.g. goal of having a healthy child) in relation to external contexts (e.g. physical state of being pregnant) and were delineated by events or by data itself. Participants envisioned reflecting on data to predict future patterns, set goals, view progress towards goals, and inform them on how to reach goals. They reported that modifying their data visualizations to explicitly model their implicit perceptions of data from BSC would be a useful aid for self-reflection. Finally, I discuss the characteristics of BSC, the importance of surfacing BSC in health data visualizations, and how subjective and implicit boundaries can be made explicit through modifying data visualizations.

5.3 Methods

5.3.1 Study Context: Pregnancy

Pregnancy is a brief period in a woman's life that is often considered disruptive. The lived experience of pregnancy is rife with uncertainty because physiological and behavioral changes that may or may not be transient [243]. During pregnancy, women may be confronted with gestational diabetes what could alter her patterns of eating, a body that disrupts her required sleep, and prenatal psychosocial stress that comes from the high-stakes rhetoric around reducing harm to the fetus [49]. For a woman having her first child, she is also transitioning into the role of motherhood, marking the beginning of further changes to routines, goals, and priorities [49]. While self-tracked data from before, during, and after pregnancy externally appear to exist in distinct phases, I examined how women post-pregnancy reflected on data from these past periods.

5.3.2 Recruitment: Post-Pregnancy Women

Participants were recruited from a pool of current participants of an ongoing clinical trial at Northwestern Lurie Children's Hospital: Wellness for Two – A Randomized Controlled Trial. The goal of the Wellness for Two study is to understand the effects of a pregnant woman's stress on the neurodevelopment of her child [247]. Participants of this clinical trial were recruited during their first or second trimester and their participation continued until the child is one year old. During their second trimester, participants of the study were asked to track their stress using an adhesive heart monitor and by answering ecological momentary assessments (EMAs) about their stress and activities for 14 weeks. Half of the participants (50 of the total 100 target sample size) in the Wellness for Two study was randomly assigned to an intervention group and the other half were assigned to the control group. For those in the intervention group, if an algorithm determined that the pregnant woman had reached a threshold of stress on one day, they would receive a wellness intervention the following day in order to reduce their stress [58]. No participant was permitted to review their self-tracked stress data in order to not influence their own stress levels. After the 14-week period of stress tracking was completed, participants of Wellness for Two were eligible for this main study.

I had IRB approval to reach out to these participants to recruit them for this study. At first contact, I explained the purpose of the study and asked them if they would be interested in participating. Participants had the opportunity to ask questions about the study before they consented. I scheduled sessions through e-mail, phone, or text and used a digital consent form. Due to COVID-19, sessions were conducted remotely over video call and screen share Recruits will were compensated $30 by PayPal for their participation.

5.3.3 Phase 1: Elicitation Interviews

The objective for Phase 1 was to better understand participants' perceptions of the uses of and interest in their pregnancy data after pregnancy.

5.3.3.1 Participants

Data saturation was reached after interviewing a total of eight participants, at which point I concluded data collection for Phase 1 [29]. Most participants had just given birth to their first child. On average, Phase 1 participants were 4.4 months post-birth at the time of the interview which resulted in a nonrepresentative sample.

Table 5. Phase 1 Participant Characteristics

	Participant #	Months since birth	First child?
PHASE ONE	1	3	Yes
	2	2	Yes
	3	4	No
	4	3	Yes
	5	7	Yes
	6	4	Yes
	7	6	No
	8	6	Yes

5.3.3.2 Data Collection

Phase 1 consisted of two parts: (1) interviews to understand general perspectives on data the participants captured (including the data from their pregnancy) and (2) an elicitation activity to understand their perception around specific types of data they collected. Each session lasted approximately one hour.

First, through the interviews, participants were asked broadly about their prior experiences with self-tracking and how they perceived their pre-pregnancy self-tracked data compared to their self-tracked data captured during their pregnancy. Next, participants were asked about their desired uses for self-tracked data that is captured during pregnancy and their uses for this data after pregnancy. The goal of this method was to surface their perspectives on the uses of data captured during a disruption and understand if and how they drew distinctions between data captured during pregnancy and data captured outside of pregnancy. More specifically, the semi-structured interview questions covered:

- Their experiences with self-tracking in general and when it comes to health (outside of the clinical trial)
    - What was tracked
    - How they decided to track
    - When started/stopped tracking
    - How the data were used at the time
    - Whether the data was shared with others, and if so, who, and why
    - How the data are used today (if it is used at all)
- Their experiences self-tracking stress during pregnancy (as part of the clinical trial)
    - What they were most interested in learning about themselves…
        - Broadly: if they could have tracked anything during their pregnancy
        - Specifically: if they had access to their Wellness For Two stress data
    - How would they use these stress data (and this use would persist after pregnancy? Why?)

After the interview portion, I conducted a textual-elicitation activity to further explore their interests when it comes to their data and ideate upon design features that may meet their needs. The goal was to identify specific examples of stress self-tracking data that the participants would be interested in seeing (e.g. Stress with Sleep and Excitement) and using these tangible examples, understand when these data are useful to them and how this "usefulness" may change after pregnancy. In sum, I first established the type of data that the participant was interested in exploring and then I probed on if and how they imagined using these data after pregnancy and when they consider the data no longer relevant or useful.

The virtual textual elicitation activity was developed using an online drawing board tool with several 'movable cards.' Each card had one type of data collected during the Wellness For Two study or other data participants may have described tracking or would have liked to track during their pregnancy. The cards were labeled as follows (see Figure 5 for screenshot):

- Heart rate
- Automatically labeled: Stressful Day vs Non-stressful Day
- Q1: In the past hour, did you experience anything stressful?
- Q2: Over the past hour, how stressed were you feeling?
- Q3: Over the past hour, did you feel you could not control important things?
- Q4: Over the past hour, did you feel confident in your ability to handle problems?
- Q5: Over the past hour, did you feel things are going your way?
- Q6: Over the past hour, did you feel difficulties piling up so you cannot overcome them?
- Q7: Over the past hour, how happy were you feeling?
- Q8: Over the past hour, how excited were you feeling?
- Q9: Over the past hour, how worried were you feeling?
- Q10: Over the past hour, how irritable/angry were you feeling?
- Q11: Over the past hour, how sad were you feeling?
- Q12: In the past hour, have you exercised or done physical activity?
- [other data participants tracked during pregnancy]
- [other data participants tracked outside of pregnancy]

- [any other desired data]

Heart rate

Automatically labeled: Stressful Day vs Non-stressful Day

Q1: In the past hour, did you **experience anything stressful**?

Q2: Over the past hour, how **stressed were you feeling**?

Q3: Over the past hour, did you feel you **could not control important things**?

Q4: Over the past hour, did you feel **confident** in your ability to handle problems?

Q5: Over the past hour, did you feel **things are going your way?**

Q6: Over the past hour, did you feel **difficulties piling up** so you cannot overcome them?

Q7: Over the past hour, how **happy** were you feeling?

Q8: Over the past hour, how **excited** were you feeling?

Q9: Over the past hour, how **worried** were you feeling?

Q10: Over the past hour, how **irritable/angry** were you feeling?

Q11: Over the past hour, how **sad** were you feeling?

Q12: In the past hour, have you **exercised** or done physical activity?

[other items tracked?]

Figure 5. Screenshot of textual-elicitation activity with movable cards that participants selected from based on interest in the data.

Participants were asked to arrange cards to create combinations of data types that they would be interested in learning about. For example, they may combine the card "Heart Rate" with the EMA question "Q12: In the past hour, have you exercised or done physical activity?" in order to investigate how often their stress overlapped with a raised heart rate with respect to occurrences of exercise as opposed to other potentially stressful events.

Participants were encouraged to think aloud as they selected cards. For each combination, I probed on why they were interested in the combination, what information they hoped to learn, and the perceived longevity of the utility of the data (*Would you want to review data from this period after your pregnancy is over?*). Due to time, I asked them to create at most three combinations. I asked them to rank which combination would be most interesting to them *during* pregnancy and which would be most interesting *after* pregnancy in order to elicit how their values changed as a result of this shift in context. Lastly, I asked how their data should appear to them once this context has shifted – if they wanted to see it at all.

5.3.3.3 Data Analysis

Due to the visual nature of the interviews, all sessions were video- and audio-recorded. Coding and refinement of interview questions were conducted iteratively throughout the data collection process. I wrote memos after each session throughout the study. After the first few interviews, I watched videos of the initial interviews, pausing to memo and transcribe data as I formed open codes. I described visual actions and representations to text as necessary [96]. Open codes were condensed into axial codes to create a code book, and the remainder of the videos were coded using the resulting code book. Through this iterative inductive approach, I identified themes. See Chapter 3 for more details on thematic analysis methods.

5.3.4 Prototype Development

Based on the data from Phase 1, I designed a prototype for use in a graphic elicitation activity in Phase 2. In this activity, Phase 2 participants used the prototypes to discuss how they wished to manipulate visualizations depicting pregnancy data according to their needs.

I simulated and graphed two years of data on Weight, Sleep, and Stress using python with numpy and the Pandas library. Graphs using the data generated were created in Excel. The Weight data was simulated based on prior studies that found that weight increased an average of 30 pounds during pregnancy and declined rapidly over the first 8 weeks after birth and slowly after that [25]. Sleep data was completely generated based on prior studies on the decline of sleep during the third trimester [97]. Given that the participants of the Wellness For Two Study only collected stress data from during the second trimester, I utilized the stress-related EMA responses from a single participant from the Wellness For Two study chosen for their adherence to answering the EMAs. This Wellness For Two participant was excluded from participating in our study in order to ensure that Wellness For Two participants were never shown their own data. I generated simulated data depicting data from before and after the pregnancy such that stress was lower during certain segments and higher during other because pilot tests found that when EMA data was generated randomly, pilot participants did not know how to interpret or respond to the data.

While each of the participants saw the same graph and data, the dates of the each of graphs were customized to match their dates of pregnancy and additional simulated data was generated so that the graph would represent data up until the month of the interview. Initially, I had chosen to simulate 3 months of data before and after pregnancy. However, after the first three participants, I learned that the 3-month postpartum period is perceived differently from the 12-month postpartum period as the mother's routine changes and the infant grows rapidly in that time frame. Therefore, I simulated more post-pregnancy data; the remaining participants were shown a total of 2-years of data: 3 months of pre-pregnancy, 9 months of pregnancy, and around 1 year of post-pregnancy. In other words, for each data type I simulated data from 3 months before their pregnancy up until the day the interview took place in order to understand the span of changes to their perceptions of data in the time elapsed. Figure 6 through Figure 8 depict the graphs used for this graphic elicitation exercise.

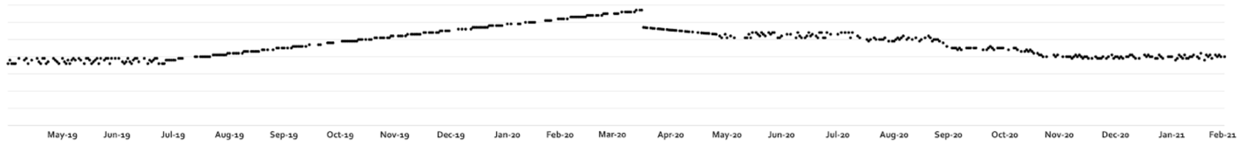


Figure 6. Graph of 2-years of generated weight data. Birth depicted mid-March 2020.

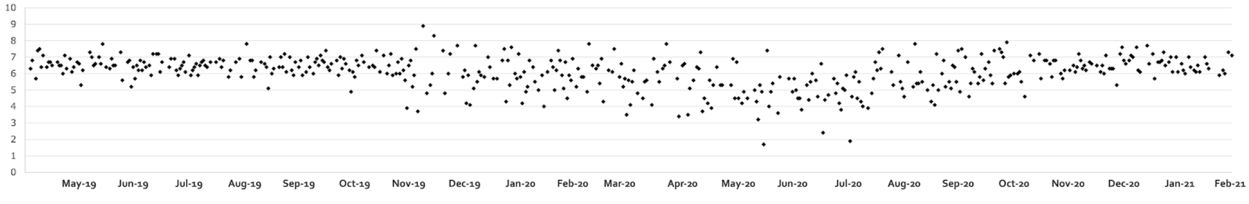


Figure 7. Graph of 2-years of generated sleep data. Birth depicted mid-March 2020.

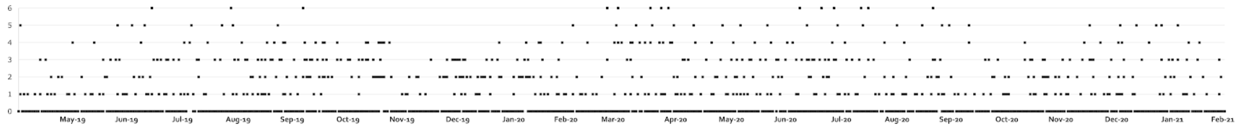


Figure 8. Graph of 2-years of generated stress EMA data. Birth depicted mid-March 2020. Note that since participants were asked to rate their stress four times a day, there are four values depicted each day.

5.3.5 Study Procedure: Phase 2: Design Evaluation and Probe

The goal of Phase 2 was to learn how the various data visualization manipulation features met the needs of the participants exploring health data from their past pregnancy and through this, better understand reasons to reflect on data after a disruption.

5.3.5.1 Participants

Data saturation was reached after interviewing a total of 13 participants, at which point I concluded data collection for Phase 2 [29]. Most participants had just given birth to their first child. On average, Phase 2 participants were 10.9 months post-birth at the time of the interview which resulted in a nonrepresentative sample. See Table 6Table 5 for a summary of participant characteristics.

Table 6. Phase 2 Participant Characteristics

	Participant #	Months since birth	First child?
PHASE TWO	9	7	Yes
	10	8	Yes
	11	10	Yes
	12	11	No
	13	12	Yes
	14	11	Yes
	15	11	No
	16	11	Yes
	17	11	No
	18	12	Yes
	19	12	Yes
	20	12.5	Yes
	21	13	Yes

5.3.5.2 Data Collection

Phase 2 consisted of two parts: (1) interviews to understand general perspectives on data the participants captured (including the data from their pregnancy) and (2) a graphic elicitation activity to explore how participants wanted to modify the three graphs (Weight, Sleep, Stress) in order to extract the information they found relevant. Each session lasted approximately one hour.

Similar to Phase 1, Phase 2 interviews began by asking the participant about their prior experiences with tracking their health and the use of this data over time. Then, participants were asked to complete the timeline activity by graphing their stress levels for the past two years while verbalizing any relevant events they were comfortable with sharing in order to allow me to ground the remainder of the interview session with an understanding of their recent life events (see Figure 9 for an example).

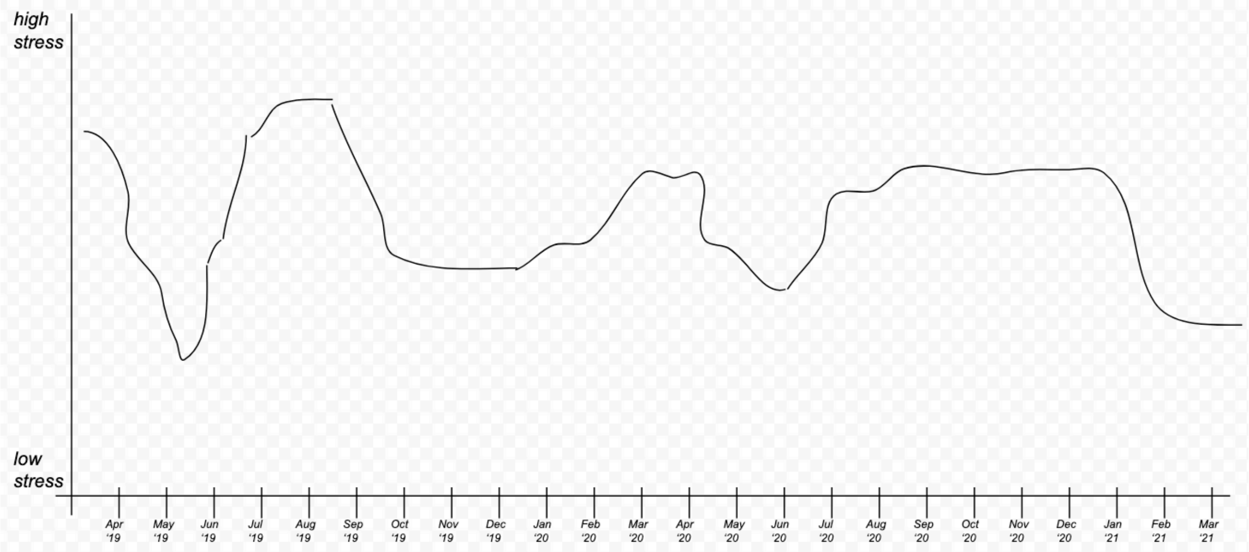


Figure 9. Completed Timeline activity

Next, participants were asked to modify the graphs of two years of simulated Weight, Sleep, and Stress data (including a 9-month pregnancy) according to what they would find relevant to their current interests. These data types were chosen based on findings from Phase 1 in which some participants perceived the influence of pregnancy on their physical health data differently from their mental health data. In contrast, when it came to weight data participants had more uniform responses: Weight data was inextricably tied to pregnancy. Therefore, I decided to use Weight data as a means of introducing the elicitation activity.

In order to prompt participants to discuss how they might want to modify the graphs, I demonstrated three potential data visualization modifications (which I named Highlight, Archive, Mark) but invited them to come up with their own suggestions if they so desired. These three initial data modification ideas were born out of Phase 1 findings, in which participants expressed how they envisioned using or not using data from their recent pregnancy. These findings suggest that participants may value the capabilities of hiding segments of data (Archive), using markers to compare segments of data (Mark), or visually distinguishing segments of data (Highlight). While sharing my screen, I then illustrated their desired changes onto the graphs and asked participants how this change influenced their interpretation of their graph. I also probed participants on the benefits and tradeoffs of modifying these graphs. Figure 10 depicts a weight graph with desired modifications as described by a participant.

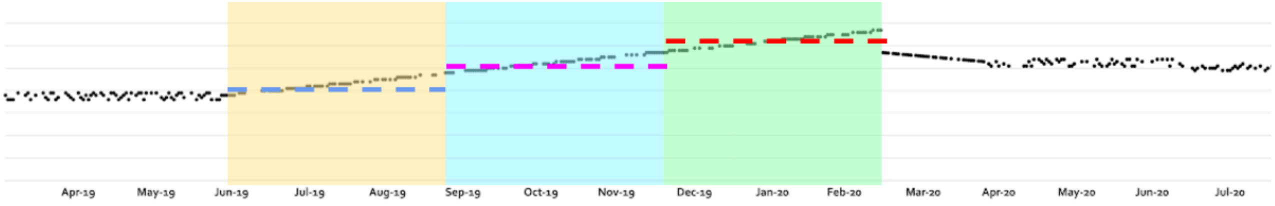


Figure 10. Sample image of modified Weight graph (P20)

5.3.5.3 Data Analysis

Similar to Phase 1, due to the visual nature of the interviews, all sessions were video- and audio-recorded. Coding and refinement of interview questions were conducted iteratively throughout the data collection process. I wrote memos after each session throughout the study. After the first few interviews, I watched videos of the initial interviews, pausing to memo and transcribe segments of data as I formulated codes. Open codes were condensed into axial codes to create a code book, and the remainder of the videos were coded using the resulting code book. Through this iterative inductive approach, I identified themes. During this process, I found some themes similar to those in Phase 1. This was expected as I asked similar questions in Phase 1 in which I had reached data saturation after 8 participants. I had chosen to ask the same questions to the participants in Phase 2 as a means of introducing the participants to the concept of reflecting on the data they had previously collected. Therefore, in the following section, I indicate which findings were drawn from both phases of the study and which findings were drawn from a single phase.

5.4 Findings

In examining participants' conceptions of and envisioned uses for data collected during a period they do not currently consider routine (pregnancy or past periods of life), I observed how participants implicitly drew boundaries as they reflected upon their data. **Here, I introduce *bounded situational context* (BSC) as a concept to refer to an individualized context that is derived by factoring internal and external contexts during a non-routine circumstance.**

Although all participants experienced pregnancy, they described differing boundaries of pregnancy with respect to their health data. When data is collected during a non-routine circumstance, internal context is critical as the individual must re-negotiate the uses of data that they otherwise found useful during their routine state. This re-negotiation was based on their goals and priorities, their understanding of their data, and as well as the external context around which their data were collected. In other words, BSCs emerged as participants discussed how they relate their impressions of the data collected to their pregnancy. While the external context of pregnancy is generally strictly a 9-month period, internal contexts did not always temporally align as the individuals' internal goals and unique lived experiences influenced the perceived boundaries of their pregnancy (see Figure 11). In order to support this sensemaking process, it is critical to understand how these data from BSCs are delineated.

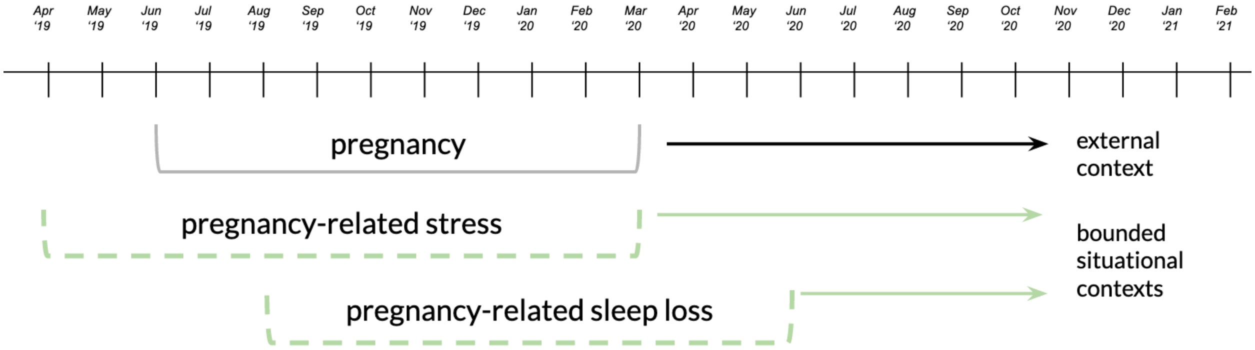


Figure 11. There may be various ways an individual can bound their data within a single external circumstance.

In the next sections, I describe how individuals defined their boundaries based on their internal perceptions, events that occurred, or the data itself. Then, I present the ways participants envisioned using or not using data captured during and around this period of disruption. Lastly, I describe how participants hoped to modify their data visualizations according to what they found relevant after pregnancy and their reception to the notion of altering their data visualizations.

The first two subsections of the Findings report themes derived from both phases of the study. The findings in the final subsection on the design of data visualizations are from analysis of Phase 2 data exclusively.

5.4.1 Creating Boundaries

As participants reflected on their data from a past period of life that is different from their current routine, they implicitly applied boundaries within their data to aid in their interpretation; they partitioned their data through what I refer to as BSCs. These boundaries of BSCs were defined by events or through the data itself and was often determined by their perception of the influence of external circumstances on their data.

5.4.1.1 Event-driven boundaries

Participants described specific events that they perceived to be delineators that separated periods of routine. These events can mark a transition to a new routine as described by P3, who had journaled for years to track her thoughts, food, stress, and migraines to detect triggers. When asked what she did with her journals from the past, she stated:

> *"Looking back at the stuff I had from years ago – I didn't want to re-read it because I don't want to relive that. So I would just shred it and recycle it and just be like that's a part of my history… It's kind of like milestones in my life, like I would re-read stuff about before being married. I'm married now, that stuff isn't relevant. Let's just get rid of stuff that's not relevant and focus on what is relevant. Then when I had a kid, I said 'Well this isn't relevant anymore, these are my old issues. Now I have a family to worry about' so let's just purge those. It's like a renewing of life and a moving on with life instead of dwelling on things."* – P3

P3's data from prior life stages or milestones were perceived to be irrelevant to her current state due to events that changed her routine or environment. She specifically mentioned two boundaries: getting married marked the end of one period and having a child marked the end of another period of life. These external events in combination with her perceptions of the type of issues she faced during this period delineated distinct periods of her life. Pregnancy was similarly referred to as a transitional period for other primigravida (first-time mothers) participants and especially in the first few months after birth, they often described struggles with the process of establishing a new routine.

5.4.1.2 Data-driven boundaries

When circumstances are ill-defined, the data itself can be used to draw boundaries between routine and periods of disruption. Many participants envisioned using data from before pregnancy (disruption) as a baseline to determine whether "normal" (routine) has resumed after giving birth. However, the boundaries of pregnancy varied depending on the individual: their values and how they perceive the type of data being analyzed. For P10, her pregnancy-related sleep was bounded at the beginning by when her sleep began to deviate from her "routine" and ended when her sleep returned to the "routine" state:

> "*Pregnancy is not representative of the rest of life. Data would be used to notice and then shelve unless I were to become pregnant again and was interested in comparing sleep patterns. I would shelve it when [sleep] starts to change until maybe 5 months post-partum might be when I pick up because theoretically it might be getting back to whatever normal is.*" – P10

When reviewing sleep data, P10 disregards data that is different from what she considers to be her norm as she believes it is not representative of her routine, which is the data she would be more interested in seeing. Unlike an event-driven boundary, the start of the disruption is not when she became pregnant, but when her sleep begins to change due to her pregnancy. The measurement she used to determine the end of this disruptive period is internally motivated by the desire to reach the point where she slept as much as she did pre-pregnancy. The disruption does not end with the event of birth, but upon the return of her "normal" number of hours slept – the data itself.

While most participants held the sentiment that pregnancy is a disruption to routine that has no relevance to their routine (non-pregnant) state, upon closer examination, I saw that whether there was value in holding on to their data from this period depended on the type of data that the individual envisioned reflecting upon.

5.4.2 Reflecting on Data from Bounded Situational Contexts

While participants described pregnancy and other past periods of life as being "different" from their current routine state, participants still thought that these data could be useful when setting goals, checking for progress towards goals, and inform them on how to reach their goals. However, the usefulness of the data was contingent upon participants' perceptions and expectations around how their prior self relates to their current or future self.

5.4.2.1 Modeling the Future

Data from the past was perceived to be useful if it was captured within a context that could reoccur. Many participants stated that weight data from their recent pregnancy would only be useful if she were to be pregnant again in the future:

> *"I would be interested in Apple watch data for the next pregnancy. I know that I had a very easy pregnancy and I was able to lose the weight fairly quickly so when I get pregnant again, I'd like to do the same thing – make sure that I at least do the same amount or around it to kind of have another easy pregnancy ideally…The data is not relevant today though. Pregnancy is just such a specific thing. Once you're not pregnant anymore, I just don't think it's relevant. How you were when you're pregnant versus how you are now I don't think is relevant."* – P18

P18's quote is emblematic of a common theme that many participants shared: physiological measurements during pregnancy are only applicable to a future pregnancy. Notably, when it came to stress data, some participants believed the stress was heavily influenced by pregnancy and could not be compared to non-pregnancy stress data while others believed that pregnancy was irrelevant when reviewing stress data.

> "*I think I could validate my hypothesis that if I exercise more, it makes me less stressed and then I sleep better...If you keep doing that, it can keep reducing stress in your life. I feel like I could use that post-pregnancy to continue to optimize the right amount of exercise or other factors that I can do to make sure that I sleep well and reduce my stress…I think this information is relevant my whole life. I just feel like this is true pregnant or not. Maybe my sleep is not going to be as good when I'm getting up five times a night to pee than when I'm not. But if I do still think that if I can show that for me, there is a correlation between exercise, stress, and sleep when I'm pregnant, the same probably holds true when I'm not pregnant.*" – P5

P5 perceived data from pregnancy to be applicable across the board while in contrast, P2 believed that hormonal changes from pregnancy influenced her stress and sleep in ways that were not applicable to non-pregnant life.

> "*I think that you're worried about different things while you're pregnant so different things are stressful to you. Like worrying about the health of the baby and what's going to happen with the birth and where are my other children going to be when I'm giving birth and who will take care of them? There are so many other things that are very specific to being pregnant that are stressful when you are pregnant and when it is over, you have the baby and that kind of stress is gone because you now have new stress of having another child and things that about that child and your life. Now I'm not worried about the stuff in the same way as I was when I was pregnant.*" – P15

P2 and P5 viewed the same data with different utility because of their different perceptions of the impact of the context of pregnancy on these data. In other words, whether data on stress was bounded by the context of pregnancy was individualized. The examples above illustrate how self-tracked health data can oscillate in and out of usefulness based on individual's perceptions of the influence of the disruption on their data. Whether the post-pregnancy context is considered applicable to future states is defined by the individual themselves.

5.4.2.2 Envisioned Post-Disruption Data Usage

After pregnancy, some participants envisioned using data from a prior routine as a baseline to measure against their new normal or to instruct them on how to return to a prior norm. Past data could be considered evidence of a prior state of health that they could eventually return to. For instance, P6 used past data as a point of reference to create personalized goals:

> *"I don't look back to half-marathon, that time, 5 years ago. I've looked back a couple years just to see what my pace was pre-pre-baby because post-baby it's really bad. But I've looked back 2018 or 2017 just to see what my pace was…It's just to see what my potential was before the baby to just see 'Hey if possible, I'd love to get back to doing that just to compare' but I know a lot of things have changed. It's a good goal for me but mentally I take a lot of that apart and say 'Well, you know, let's factor all these other things too.' It gives me a goal to look back to and say, 'I was able to do this once, maybe I can do it again.'"* – P6

This participant acknowledged that her life circumstances are different now after having a baby, but her running data from a few years ago could still be used to set a goal. However, choosing more recent data (3 years as opposed to 5 years) as a baseline can also feel more accessible as it is anchored in what they have more recently achieved.

Data from past achievements could also be used to inform the participant of behaviors that led to their previous success. Several participants recalled their food tracking experience from a period in which they were trying to lose weight (e.g. a wedding) and envisioned the same data being useful post-pregnancy, a period in which they had the same goal: trying to lose weight.

> *"We were getting married so that, I actually tracked calories for 2-3 months. I was pretty regimented about it. I was able to lose about 10-15 lbs. It was good…I want to know how long it actually took and how much specifically I was able to lose…my activity level, too – how active I was. Around that time I was playing a lot of soccer and I was probably running pretty well to keep up with that."* – P6

P6 saw value in using data from this prior routine for its potential to be instructive for her current weight loss plans. Similarly, P14 looked to a period where she felt she was at her healthiest and thought of the ways the data could inform her activities to help her reach the same goal today.

> *"Data from 6 years ago would still be useful because I feel like I was my healthiest then. It would be nice to see the things I was doing right or the things that were helpful, or 'working'…Maybe it would tell me times that were a struggle for me, like this time of day when I eat the worst, so I need to be mindful of that time of the day. Or these meals that were really healthy and worked for me, or I must have liked this activity because I did it a lot."* – P14

Under new circumstances – after having a child – some participants looked to prior routines with the expectation that their life post-disruption would be similar and therefore they could adopt activities that had previously led to success.

On the other hand, when it came to stress data, due to the multiplicity of potential causes for stress that are not captured and the immediacy of stress, some participants felt that stress data may not be informative for long after the data was collected:

> *"It's all backward looking, so I can't do anything about it. Like if I knew in the moment 'Oh my heart rate is high' or maybe I'm more stressed than I realized that I'm paying attention to or something like that. I think it could have been useful in the moment to try to do something in response to that to try and lower my heart rate or do something like that."* – P16

In contrast to weight data, where the reasons for weight change were usually more clearly connected to eating and exercise, stress data was not overtly actionable after the stressful period. As a result, stress data became less useful over time.

5.4.3 Modifying Data Visualizations

The BSCs that participants formulated were dependent on their perception of pregnancy in relation to the particular type of data they were asked to reflect upon. This section describes the various ways in which participants hoped to modify data visualizations according to BSCs to aid in their self-reflection and their desired use or non-use of past data.

5.4.3.1 Data visualization-Modifications

Archive. By archiving data, participants were choosing to hide segments of data from view. This feature was desirable when the data are perceived to not be useful to reflect upon. Oftentimes these data were not useful because (a) the data were captured from a period when the participant felt that they were unable to act upon the data or (b) the data are not relevant to the participant when they are not pregnant.

Most participants decided to archive all 9 months of their weight data from pregnancy. This was primarily due to a lack of control over their weight during pregnancy and the inability to exercise for a period after birth. P2 envisioned using data from before pregnancy to answer the question: *"Is my current weight back to pre-pregnancy weight yet?"* This participant then determined that once she reached her pre-pregnancy weight, the data from pregnancy would no longer be relevant to her because she had reached her goal. In this example, the *Archive* feature would be useful to hide her pregnancy-related weight data based on her individual boundary.

Similarly, for sleep data, participants wanted to hide data from the period when they could not sleep well for reasons outside their control. For both these circumstances, the boundaries were determined by when they felt that they could exercise again or when their infant slept through a full night. Figure 12 shows the graph of a participant that wished to hide all "not normal" sleep data.

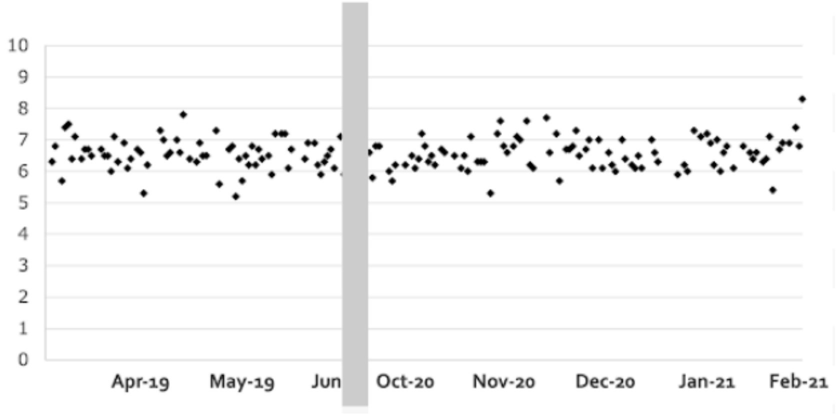


Figure 12. P8 Archived all sleep data that looked "not normal."

Whether or not participants archived stress data was dependent upon how the participant perceived pregnancy as a factor that impacted their stress levels. For individuals who stated stress was not impacted by pregnancy, they treated all stress data as 'routine' and did not seek to modify their stress data visualization at all; there was no stress-related pregnancy BSC. For those that believed that the hormonal changes were a core part of being pregnant and influenced their self-reported measures of stress, stress data during the 9-month period was either highlighted or archived. P14 considered stress data from *before* pregnancy to be within the boundaries of pregnancy itself due to fertility-related struggles and stress and therefore, wanted to archive a 10-month period.

Mark. The concept of marking data based on segmented periods such as BSCs was found to be useful in particular for tracking progress towards an established goal. A 'baseline' was a reference point typically calculated from periods that represent a high point (e.g. a period when one felt most physically fit), low point (e.g. a period of known physical inactivity), or a norm (e.g. a period categorized as routine). Figure 13 illustrates how a participant first used the Archive feature to remove the 9-month pregnancy period and Mark feature visually convey her progress on her journey to getting back to her pre-pregnancy weight which she considered to be her normal weight.

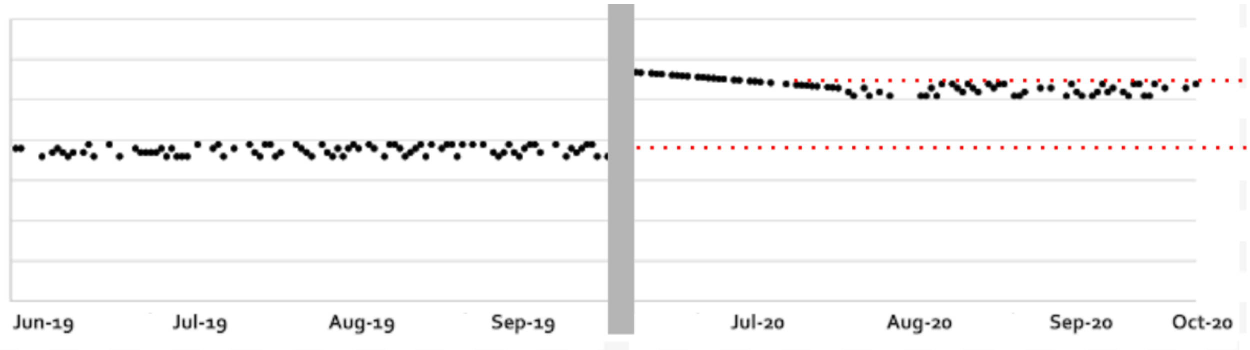


Figure 13. P9 Archived weight data during the 9 months of pregnancy and marked pre-pregnancy weight and current weight.

Figure 14 depicts how P20 wanted to modify her sleep data visualizations. She felt that her sleep patterns changed rapidly after birth. Therefore, she wanted to mark the average hours of sleep she got within 1-month periods as her child's sleep changed month by month so that she can observe trends. Furthermore, these boundaries would ensure that the "average hours of sleep" calculation encompassed only the last month's data which is most recent and therefore the most relevant to her. Overall, the mark feature was useful to summarize data that was bounded within certain periods marked by routine or disruptions to routine – individualized BSCs.

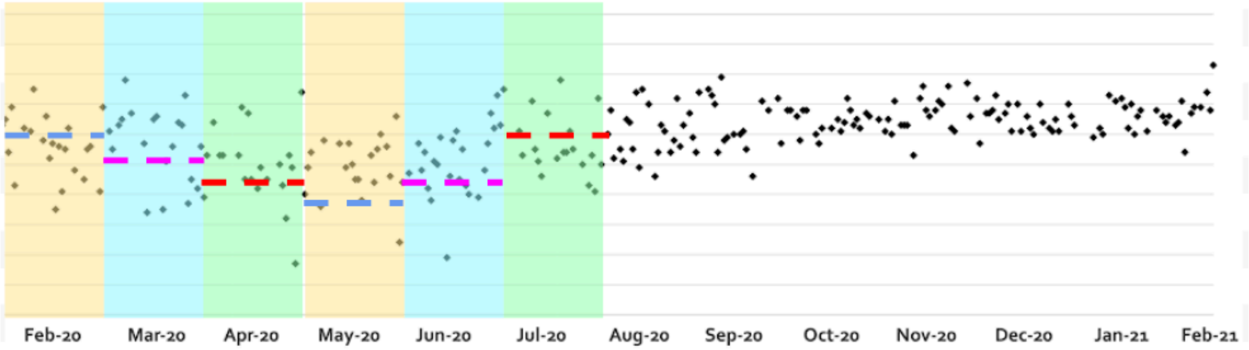


Figure 14. P20 Bounded her data month by month for the first few months after birth until her sleep (and her baby's sleep) stabilized.

Highlight. The highlight feature was often used with text annotations to label events that may have influenced the data. Figure 15 depicts P12's weight data visualization in which she used the Highlight feature to label events that she believes influenced her weight data: the purchase of exercise equipment and moving to Florida where she was able to escape the cold of Chicago and spend more time outside. The need for text annotation or the Mark feature to accompany the highlighting suggests that visual distinction of periods are not useful on their own unless there are descriptive reminders of why the segment of data is highlighted or calculations that summarize (e.g. average) the enclosed data. A few participants illustrated overlapping highlights, depicting overlapping contexts; understanding how to design for overlapping contexts will require more research.

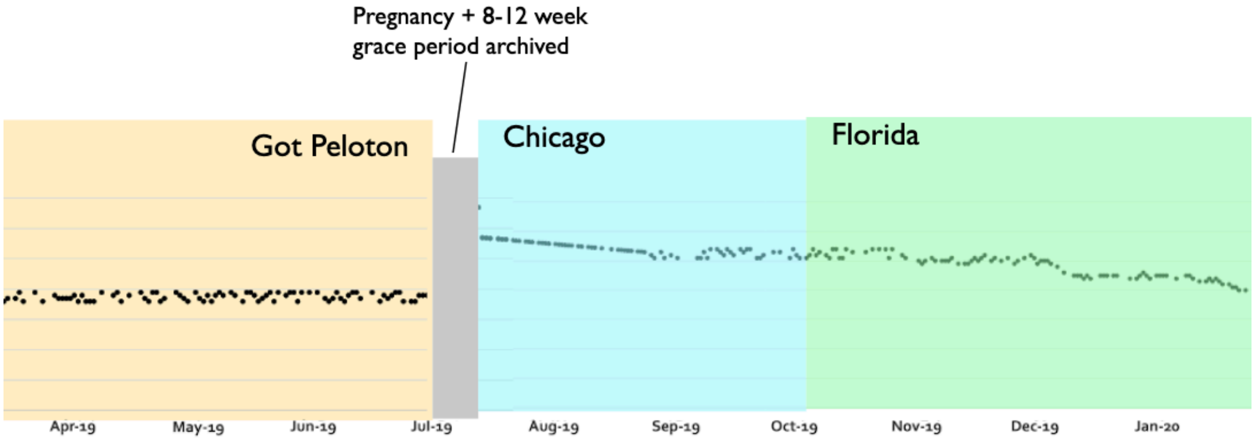


Figure 15. P12 Bounded her weight data by events that she felt influenced her weight.

5.4.3.2 *Attitudes towards modifying data visualizations*

The participants all saw value in the ability to modify graphs of their health data. Highlighting segments of data, archiving irrelevant periods, and marking goals and progress visually allowed participants to focus only on the data that is most useful to reflect upon:

> *"I find it easier to look at a graph that I've highlighted to remember that this was the period. So I'm not going back and checking 'What was this again?' I've already identified the period I want to be looking at…It'll trigger in my mind 'Oh yeah there was this other stuff going on. This is where I was at that time. It was during a certain period.'"* – P16

Through highlighting specific periods of disruption, P16 found that the data visualization-modification features can be helpful to aid in recall and sensemaking. Furthermore, when it comes to periods of disruption, the ability to hide data can be particularly useful as some participants mentioned that data from certain periods of the past could be distressing in addition to being irrelevant to their current state.

The primary barrier to the creation of these modified data visualizations is that participants imagined it would be a burdensome manual task. Participants mentioned that they barely have the time to track and review their data as they care for a new child; labeling their own data would be unlikely. While annotating boundaries of BSCs may be less time consuming than labeling individual data points, passive methods of surfacing potential boundaries can help expedite the task of modifying data visualizations.

5.5 Discussion

This study investigates how health data tracked during non-routine circumstances can be used by individuals once that period has passed. To help conceptualize the boundaries that individuals implicitly drew around their data, I introduced the notion of *bounded situational contexts* (BSCs) to encompass the ways that individuals take internal and external contexts into account when assigning relevance to segments of their data. Moreover, I explored how visualizations of data collected during BSCs can be modified to support the lived experiences of the individual. In this section, I outline the specific characteristics of BSCs, discuss how the framing of BSCs can support sensemaking during the evolving lived experiences of self-trackers, and detail opportunities for design to move forward in this space.

5.5.1 Characteristics of Bounded Situational Contexts

In *Studying Context: A Comparison of Activity Theory, Situated Action Models, and Distributed Cognition*, Nardi argues that context as a whole is comprised of the fusing of external and internal contexts [178]. External contexts can be physical constraints (e.g. environments, tools) or social influences (e.g. cultural norms, obligations) while internal contexts can be personal goals, skills, and preferences. With respect to the practice of self-tracking and interpreting the resulting data, an understanding of what constitutes routine can serve as a consistent frame of reference to situate anomalous events and the resulting data. When a routine is disrupted, external contexts may be unfamiliar; therefore, the individual must rely on internal context in order to interpret the situation and the data [178]. Results from this study illustrated how participants determined the boundaries of a disruptive period, pregnancy, in various ways depending on their own experiences of external and internal contexts. For instance, an individual may consider certain types of data within disruptive periods (i.e. weight during pregnancy) to be distinct from routine life while other types of data (i.e. stress during pregnancy) captured during the external context of a disruption is not distinguished from routine. How these determinations were made was based on how participants integrated internal and external contexts as they interpreted their data.

Bounded situational contexts (BSC) refers to a personalized context that an individual implicitly defines by factoring internal and external contexts during a non-routine circumstance [209]. BSCs have two primary characteristics: (1) they are not inherently recurring (2) individuals seek to segment data captured during BSCs from other data in their lives. Below, I detail each of these characteristics and highlight prior literature in HCI other than pregnancy in which this concept may serve to aid in the extraction of insights to support sensemaking through design.

5.5.1.1 *Non-Cyclical Circumstance*

Among the body of work on the use of context to interpret self-tracked data, most have focused on individuals that relied on the cyclical nature of recurring contexts when interpreting data and determining subsequent actions. For instance, the notion of *contextual frames* refers to a subset of external and internal contexts that model routine or cyclical factors that shape health behaviors and subsequently health data [193]. Central to contextual frames is the recurring nature of contexts and therefore a consistency in how data such as blood glucose data should be leveraged for diabetes management. However, studies on other chronic conditions that often do not follow regular patterns or continuously evolve over time have found that individuals rely on internal contexts when external contexts are unfamiliar or when there is a dearth of personalized health information [155].

BSCs contextualize data captured during circumstances that are not necessarily recurring or cyclical. These findings highlight the complexities around the reflection of data from a period that is considered disruptive, transitional, and not guaranteed to recur. Prior works in HCI have studied individuals practice self-tracking during events or moments that are non-cyclical or non-recurring such as progressive chronic illness such as Parkinson’s Disease, spinal cord injury, multiple sclerosis, fibromyalgia, periods of self-experimentation, and menstruation for transgender men and individuals with endometriosis [26, 70, 120, 155, 170, 177, 180, 242]. These are periods in which aspects of health are irregular or not well-understood by the individual. These studies have discussed how the changing bodies and disease states require individuals to build new knowledge in order to better understand their bodies in the process of sensemaking and subsequently adjust their self-care practices [70, 242].

Tracked data is interpreted with respect to its context; given that some contexts may never arise again, it is possible that these data captured during this period would not be useful to integrate into a broader reflective practice. The findings of this study saw that most participants referred to weight data from pregnancy as an “outlier” with “no impact on life” unless they were to become pregnant again. The data was perceived to reflect an outlying period in life wherein an individual cannot form patterns, compare them to other periods of life, or derive insights that can be used to inform actions outside of the specific context during which they were collected. Other studies that explore the lived experience of a progressive chronic illness have similarly considered how data collected may not be useful in a future advanced disease state:

> *"[Parkinson's Disease (PD)] symptoms vary a great deal over time and progression can be difficult to predict...researchers must consider when and how to present this information to users. Since framing of self-tracking data can influence users' self-perception, presentation of information about disease progression must be thoughtfully designed so as not to foster a sense of helplessness and depression" [170].*

More explicitly, based on their findings of a study on how menstruators reflected on their menstrual tracking data, Fox et al. argue for:

> "*…expanded forms of bodily sensemaking, ones that might emphasize the multiplicity of practices conducted to align with people's goals and identities—over the models that come with algorithmic ways of knowing—and dimensionality, rather than a user's relation to averages or norms*" [70].

In other words, data collected during BSCs may no longer be of use to the individual in their self-reflection and instead can be misleading or detrimental when viewed wholesale or through indiscriminate aggregation across different dimensions of data and periods of life. Still, there can be potential value in preserving data captured from BSC including cases in which such circumstances arise again.

In sum, BSCs are characterized by the infrequency or non-recurring nature of the circumstances surrounding when the data were captured. Consequently, data captured during BSCs may only be compared to data captured during the similar circumstances and should likely not be aggregated with other data.

5.5.1.2 *Data Distinction and Need for Curation*

BSCs are characterized by an individual's need to distinguish and curate data from specific non-recurring periods in relation to data outside of this period. Elsden et al. have previously highlighted the need for designers to support individuals' subjective interpretation of their data through data curation [56]. Previous work has found that selecting subsets of data or highlighting specific events can support meaning-making and storytelling [60, 82, 88, 92]. Individuals reflecting on past data have been observed relying on "key events" in order to derive meaning from their data and form a narrative more so than chronology [56]. In an exploration of *visual cuts* of self-tracked data, Epstein et al. suggest that highlighting anomalies from routine can support changes to routine activities itself [60]. In the previous section, I defined BSC as data captured during non-recurring contexts; individuals' desire to curate data to draw a distinction between routine life and the anomalous contexts is a natural extension to this initial premise and therefore the second characteristic of BSCs.

Defining specific segments of data as a BSC is the practice of disambiguating the segments of data that fall within the boundaries of what individuals consider relevant to their self-reflection. Designers should consider how individuals determine how they would want to curate data from their pregnancy-based BSCs as opposed to data captured outside of this period. The perception of BCSs and their distinction from other periods of data is highly individualized and closely tied to the data stream being reflected upon. Though participants readily acknowledged the explicit, external context of pregnancy as a 9-month period, internally, they drew different boundaries in how they perceived specific data captured. Boundaries were often determined by the individuals' perception of causality. For instance, P14 bounded the beginning of her pregnancy-related stress data starting before pregnancy due to her concerns about whether her IVF procedure would be successful and designated the end at the birth of her child. P10 defined the boundaries of her sleep disruption based not on the external context of pregnancy, but how she slept as a pregnant woman and as a parent caring for a newborn. These examples illustrate how depending on the type of data the individual is interpreting (e.g. weight, sleep, stress), participants assessed different causal factors to determine whether the data continues to be relevant to them or whether the utility of data is confined to the specific outlying context. Future designs can facilitate sensemaking not only through labeling context, but by applying their understanding of how individuals perceive specific data streams throughout various contexts.

Over the course of life as disruptive circumstances occur or as the trajectory of illness advances to new stages [177], the framing of BSC can be used to refer to how individuals mark the end of one period and how it is distinct from the next. While not all external circumstances may be as clearly defined as pregnancy, this study provides insight to how individuals desire to curate their data: through individual perceptions of causality and preconceptions of how specific data streams are influenced by contexts that occur. In turn, this research provides designers a lens through which to develop data visualizations for data collected during BSCs.

5.5.2 Understanding BSC in Long-Term Self-Tracking

Adopting Elsden et al.'s definition of *data-work* situates the task of defining BSCs as an essential step in the process of sensemaking. The term data-work describes "how individuals interact with personal data in situ" [56]. HCI scholars have previously described how the process of sensemaking involves continuously creating associations between aspects of daily living and the data collected [56, 157, 242]. They described a *negotiation* between what one knows and what data conveys in order to establish what is considered relevant to the current situation. Technology that embraces BSC as a lens can facilitate sensemaking by aiding users in classifying segments of data that should be treated differently for a variety of reasons including disease progression or discomfort in seeing data from certain periods of life.

Data work involves evaluating and reevaluating what continues to be relevant over the course of time. This evaluation occurs through a negotiation of personal goals (internal context) and environmental constraints/lack of control (external context). In the age of personalized healthcare, more individuals will gain access to data that were captured during disruption or transition. The data work observed in this study involved participants weighing which data captured during and before pregnancy continue to be relevant to them around a year after pregnancy. Due to the timing of this study, most participants were reflecting their perceptions of their data under the circumstances of giving birth and caring for a newborn during the beginning of a pandemic. When asked how they reflect on their physical activity data, several described not being as interested in their data as they would be during "normal" circumstances because they were busy caring for the child and trying to avoid infection. By exclusively focusing on external contexts (e.g. considering pregnancy as a static 9-month period or a pandemic), we oversimplify individual's implicit data interpretation practices [153]. Beyond the explicit 9-month period or the ongoing pandemic, it is important to recognize how participants' values has shifted from physical activity to caring for a child and trying to keep her child safe. Therefore, when a participant determines that a certain segment of data captured is not useful for her self-reflection, this boundary is not determined solely based on the 9-month pregnancy period nor strictly by the pandemic. During overlapping disruptions, it is imperative to understand the underlying BSCs when interpreting data.

As more data are continuously captured, the likelihood that data is inadvertently collected during a disruption or period of transition will increase. The framing of BSC will be progressively important because these data will need to be distinguished and evaluated separately from what is 'routine.' As opposed to short-term tracking, which has been defined as tracking on the scale of weeks or months, long-term tracking (tracking on the scale of years) inevitably captures data across different routines as life events occur [166]. Technology is designed to present our past histories to ourselves through data, social media posts, and other digital artifacts [56, 88, 258]. Currently, in the dominant commercial self-tracking technologies, there are few built-in capabilities that would allow users to indicate which data are useful or control how data is displayed. In the next section, I discuss ways that technology can be designed to support the implicit work and values that individuals impart as they reflect upon their data from BSCs.

5.5.3 Design Considerations for Bounded Situational Contexts

Currently, data that is collected is designed to exist in perpetuity. Findings reported from this study illustrate how tools that leverage past data to determine current health status or proposed interventions should carefully consider how data borne out of BSCs should be represented in data visualizations as these data may be unapplicable to the individuals' future state. This study is an early work investigating how BSCs can be represented through data visualization-modifications. The limited work in this area is unsurprising given the appeal of personalized healthcare which promises higher prediction accuracy in exchange for more data to train algorithms – as opposed to the relinquishing of data that are subjectively considered irrelevant [215]. Below, I apply prior research on personal informatics systems that aim to present supplementary context in ways that can support the reflection of data from BSC. First, I discuss how individuals may establish boundaries on their own data based on perceived periods of disruption by building on prior research on semi-automated tracking methods. Then, I describe how these segments of data can be modified to support self-reflection.

5.5.3.1 Semi-Automated Development of Personalized Boundaries

There is a fundamental misalignment between the ethos of modern-day self-tracking culture and the lived experience of self-trackers. Through this exploration of BSC, I identified how the use of data change during and after life disruptions which reveals the latent assumption that self-tracking data can be used in perpetuity towards self-optimization [150]. The practice of health management involves periods of disruption and recovery from disruption in which data (depending on when it is collected) are held in different regard. For instance, there were participants who were interested in their weight data during routine periods yet were no longer interested during pregnancy as well as participants who were exclusively interested in weight data during pregnancy but not during routine periods. In another example, participants reflected on low hours of sleep soon after giving birth (in which they have little to no control over the situation) differently than low hours of sleep during routine periods of life (in which they may try to extract causes and take preventative measures). As opposed to focusing purely on cyclical or recurring patterns, designing visualizations that take BSCs into account poses a unique challenge in that each new disruption may require the individual to re-negotiate which collected data retains its utility and which may lose utility during these periods. Surfacing BSC would enable self-reflection of data in terms that the individual may better understand, but given its subjective nature, it would be a highly burdensome task.

One way of reducing the effort required to label periods in data is through the use of *semi-automated tracking*. Developed in response to the need for a low burden means of collecting self-tracked data, semi-automated tracking combines manual and automated data collection methods. The goal of such systems is to lower the burden on collecting data that may not feasibly be collected purely automatically while also promoting awareness of behaviors [33]. In an example of a semi-automated tracking system that collects multiple data streams, *MoodRhythm* continuously captures behavioral trends such as sleep through a smartphone while requiring that the user answer questions about their mood [246]. These data are later processed and insights that integrate the multiple data streams are presented to the user. Other examples of technology that employs semi-automated tracking techniques focus around combining "smart" automated detection to scaffold manual tracking efforts. For instance, *MyBehavior* requires manual food logging, but a wrist-mounted wearable for eating detection can remind the wearer to log their food at appropriate times [192, 233]. Semi-automated tracking systems have been posed as a way to aid self-tracking during evolving and complex situations such as chronic disease management. It can also play a central role in the creation of data visualizations that reflect BSCs.

BSCs are derived from external contexts and internal contexts. External contexts can be detected automatically while internal contexts can be manually labeled with the aid of automation, perhaps triggered by changes in external context. For instance, Choe et al. have previously suggested:

> "*Inferring context using other data streams (e.g., capturing "gym" by combining location and activity) or importing them directly from other apps (e.g., Foursquare) can reduce the capture burden*" [34].

Of note, these examples are all external contexts. With respect to this study, an individual can label which streams of data they feel could be affected the non-routine period of life, pregnancy. For example, a consistently elevated heart rate can be automatically detected and may indicate pregnancy[8]. A user could then be alerted and have the ability to manually label the start of a BSC within the specified data stream. Once BSCs are labeled, individuals should

---

[8] https://www.reddit.com/r/fitbit/comments/445ppj/hr_reading_consistently_high_last_few_days/

have the ability to modify how these periods of data are integrated into their broader practices of self-reflection and data visualizations – if they are to be included at all.

5.5.3.2 Varying usefulness of data

Once individuals have labeled boundaries, designs should enable them focus only on the streams of data they consider to be relevant. As individuals define the boundaries of BSC, they also determine which data is pertinent to their understanding of their health and which data should be excluded from their analysis. Yet, much research on personal informatics systems in HCI visualize data with a static level of prominence to the individual reviewing their data [34, 141]. Furthermore, while early work has found that it is common for individuals to pause or stop tracking intermittently over the course of years [166], few studies have considered the value of deliberate pruning of data that are no longer informative to the individual. These early works have similarly explored the theme of diminishing returns as data is tracked over the course of a lifetime or a disease state [144, 155, 170].

The data visualization-modification features evaluated in this study: Highlight, Archive, Mark, conveyed implicit boundaries explicitly. Allowing participants to visually draw visual boundaries in their data enabled them to process and interpret their data more quickly. In this study, participants were able to annotate entire periods of data with potential factors that they perceived to influence their data (e.g. highlighting then annotating a trip to Florida in the winter to answer the question of whether her weight decreased as a result of being able to go outside). The "Mark" feature enabled participants to see metrics calculated within personalized boundaries to surface insights (e.g. Average number of hours of sleep for each trimester) or set benchmarks (e.g. Setting goal weight based on weight before pregnancy). The use of annotations and contextual information to trigger reminders align with findings from prior work on the use of context to facilitate sensemaking of self-tracked data [34, 196, 235] . Finally, the "Archive" feature, which allowed individuals to remove segments of unwanted data entirely gave the individual capabilities to leave out entire streams of data from view. During BSCs, data may be not useful or at worst, harmful; hiding the data from view can prevent rumination [55, 170]. This study highlights the dynamic relationship between human and data which can be explored through future studies that examine the journey of aging and how individuals perceive and use data across their lifetime.

While this study only explored manual annotation of causal features, calculations, and removal of data between segments, there are further opportunities for design to support sensemaking once these personalized boundaries are made explicit. For instance, individuals may seek to compare of segments of data (e.g. comparing sleep data across multiple pregnancies). In sum, data visualization systems that display self-tracked data should be designed to allow individuals to designate their perceived boundaries through semi-automation and enable them to implement modifications to their data visualizations accordingly.

6 CONCLUSION

In this chapter, I summarize the contributions of this dissertation research and discuss the limitations of and future directions for this work.

6.1 Research Questions: Preliminary Study

- (1) How can self-tracked data be used in mental health clinical care?
- (1a) What are mental health care providers' perceptions of self-tracked data?
- (1b) What are the benefits and challenges of using self-tracked data in mental health clinical care settings?

Innovations in passive tracking technologies offer an unprecedented channel of remote patient monitoring. This increase in availability of commercial self-trackers is occurring at a time when there is a cultural shift towards measurement-based care in mental health treatment. While the majority of HCI studies on self-tracking technology and health focus on physical health, there are fewer works on the opportunities for digital tracking technologies to advance mental health treatment. To explore the opportunities that these technologies can offer in clinical treatment processes, it is imperative to understand mental health care practitioners' perspectives on these new tools and concerns around their adoption.

We conducted an empirical investigation on the perspective of mental health care providers on the use of sensor-based patient-generated data (sPGD) in their work. The result is a descriptive contribution expanding the body of literature on sPGD and mental health. We learned that sPGD would provide more value if it could be correlated with standard clinical measures. Furthermore, participants envisioned ways in which discussing sPGD could support talk-based treatment practices. Despite these potential benefits, aligning the use of a new stream of data with trusted clinical metrics would take more research to ensure the preservation of patient safety, time, and resources – all of which are limited.

6.2 Research Questions: Main Study

- (2a) How do self-trackers define which data captured during non-routine circumstances are useful to reflect upon?
- (2b) How do self-trackers envision reflecting on data after the non-routine period has passed?
- (3a) How should we think about designing data visualization-modification features that support the reflection of data after the non-routine period has passed?
- (3b) What are the benefits and challenges of modifying data visualizations that depict a non-routine period?

Studies on the use of self-tracking tools and the interpretation of self-tracked data often assume that data is captured during a routine context and technology is subsequently designed under the same assumption. As illustrated by the preliminary study, a non-routine circumstance in which data is collected may influence the interpretation of data. More specifically, in the preliminary study, a clinician questioned whether the health data collected while a patient is undergoing treatment is representative of that patient's health once they return to their routine. Disruptions and transitions that break an individual's routine are common. Yet there has been little research on how individuals reflect on data after such a disruption has occurred and how they determine whether data captured during an outlying period remains useful to reflect upon.

I conducted an empirical study to examine how women envisioned reflecting on their past self-tracked data with respect to a recent pregnancy. Through observing how participants perceived their pregnancy with respect to their non-pregnancy data, I found that they implicitly drew boundaries between these periods based on events that occurred or through the data itself. Although all participants experienced pregnancy, they had varied perceptions of the boundaries of pregnancy as it related to their health data. To explain this concept, I introduced the notion of bounded situational contexts. A bounded situational context (BSC) refers to an individualized context that an individual defines through factoring internal (e.g. personal goals, background) and external (e.g. physical or social) contexts during non-routine circumstances. This conceptual contribution can be used to explain the temporal misalignment between the external disruptive period and the period of data that individuals internally perceived to be related to the disruption.

After the non-routine period has passed, depending on the individual's perceived relationship between the stream of data being assessed and the context of pregnancy, participants either imagined only reflecting on their data if they were to become pregnant again or imagined reflecting on their data as if it were captured during a routine period of life. All participants believed that data on their physical health that was collected during pregnancy was no longer relevant after they had given birth. The perception of stress data, in contrast, depended on whether the individuals felt that their stress was affected by pregnancy; if they believed that stress was not affected by pregnancy, then stress data captured during pregnancy was undifferentiated from stress data during routine life. These determinations were highly subjective and individualized. Yet, most commercial self-tracked data visualizations exhibit an implicit assumption of a consistent baseline. Data that is collected is factored into the individuals' health assessment in indiscriminately and in perpetuity. These designs fail to visibly take into account BSCs; by giving the individual tools to modify data visualizations of their health data, we can allow for visualizations that reflect individuals' desired use or non-use of data from periods of disruption.

Design contributions of this dissertation include the development of data visualization-modification features which can be used to reflect BSCs and the empirical evaluation of their perceived utility. Participants found value in the ability to (a) Mark graphs based on calculations that would allow them to summarize and compare data from various BSCs (b) Highlight segments of data to visually bracket BSCs to support sensemaking, and (c) Archive data from BSCs when it is no longer relevant. Participants found that these capabilities would allow them to focus on the data of interest more quickly and glean information that more accurately reflects their health through the removal of data they deemed irrelevant. The challenge remains as to how to enable individuals to modify their data according to BSCs with low burden, but prior research on semi-automated tracking provide a foundation upon which to carry forth this work. These studies are an early step towards the goal of enabling self-trackers to take control over their own data throughout various life processes.

6.3 Study Limitations and Future Work

Bounded situational contexts are formulated in response to the non-routine circumstances or disruption that has occurred. In this dissertation, I explored two different disruptions: an intensive outpatient treatment program and pregnancy. Different disruptions have various properties that can influence how individuals perceive their corresponding BSCs. For instance, in the preliminary study, externally, there is a clear line drawn between when the patients were in treatment and when they returned home. In the main study, participants understood that externally, pregnancy ended at the birth of their new child. Unlike the non-routine circumstances that were investigated in this dissertation, a disruption to routine such as a natural disaster may not have as clearly defined external boundaries or imagined impacts on data. Future research should investigate how individuals conceive of BSCs derived from disruptions that are not as directly perceivable.

Over time, individual's reflections on their data and themselves may change [122]. The main study interview took place during a 1-hour session months after the participant had given birth. While this study had identified how women perceived the boundaries of pregnancy data around one year after birth, these boundaries may change if they were to review their data during a subsequent pregnancy. As individuals progress through life, gain new experiences, and age, their perceptions of past events may also change [149]. Indeed, developers designing algorithms have struggled to depict past moments as individual's recollection of their own histories are often in flux [122]. A future longitudinal study would allow us to better understand how data from the past are used as new events occur and furthermore, how BSCs are molded and evolve over time.

Finally, given the possibility of corporations and institutions analyzing a lifetime of data throughout various contexts and settings, the use of long-term tracked data has ethical and social implications. In the era of digital health, while the "democratization" of medicine allows patients to have access to detailed measures of themselves instead of relying on more paternalistic models of care, often these same data are used for surveillance [218, 238]. Critics of digital health tracking envision that future of tracking practices will be an obligation imposed by employers or insurers. When individuals do not have the ability to modify their own data, the use of data from past stages of life, diseases that have since passed, or stressful periods in history may be misleading or reinforce existing structural inequalities [147, 148]. With the growing number of companies seeking to mine big data for predictions, researchers and designers alike should develop systems and technology to protect the individuals that seek to provide only the information relevant to their needs. This dissertation is part of a growing body of work that expands our understanding of how to create tools that can enable individuals to control their data, craft their histories, and tell their stories.

7 REFERENCES

[1] Gregory D Abowd, Anind K Dey, Peter J Brown, Nigel Davies, Mark Smith and Pete Steggles. 1999. Towards a better understanding of context and context-awareness. In *International symposium on handheld and ubiquitous computing*, Springer, 304-307.

[2] Ana M Abrantes, Claire E Blevins, Cynthia L Battle, Jennifer P Read, Alan L Gordon and Michael D Stein. 2017. Developing a Fitbit-supported lifestyle physical activity intervention for depressed alcohol dependent women. *Journal of substance abuse treatment*, 80. 88-97.

[3] Anne Adams, Peter Lunt and Paul Cairns. 2008. A qualititative approach to HCI research.

[4] Neera Ahuja, Errol Ozdalga and Alistair Aaronson. 2017. Integrating mobile fitness trackers into the practice of medicine. *American journal of lifestyle medicine*, 11 (1). 77-79.

[5] Minna Aittasalo, Seppo Miilunpalo, Katriina Kukkonen-Harjula and Matti Pasanen. 2006. A randomized intervention of physical activity promotion and patient self-monitoring in primary health care. *Preventive Medicine*, 42 (1). 40-46.

[6] Ali Akbari, Reese Grimsley and Roozbeh Jafari. 2021. Data-driven Context Detection Leveraging Passively Sensed Nearables for Recognizing Complex Activities of Daily Living. *ACM Transactions on Computing for Healthcare*, 2 (2). 1-22.

[7] Astier M Almedom. 2005. Social capital and mental health: An interdisciplinary review of primary evidence. *Social science & medicine*, 61 (5). 943-964.

[8] Adrienne H Andrew, Gaetano Borriello and James Fogarty. 2013. Simplifying mobile phone food diaries. In *2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*, IEEE, 260-263.

[9] American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

[10] Amid Ayobi, Tobias Sonne, Paul Marshall and Anna L Cox. 2018. Flexible and Mindful Self-Tracking: Design Implications from Paper Bullet Journals. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1-14.

[11] Anna Bagnoli. 2009. Beyond the standard interview: The use of graphic elicitation and arts-based methods. *Qualitative research*, 9 (5). 547-570.

[12] Nikola Banovic, Tofi Buzali, Fanny Chevalier, Jennifer Mankoff and Anind K Dey. 2016. Modeling and understanding human routine behavior. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 248-260.

[13] Jakob E Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg and Lars Vedel Kessing. 2013. Designing mobile health technology for bipolar disorder: a field trial of the monarca system. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 2627-2636.

[14] Ian Barnett, John Torous, Patrick Staples, Luis Sandoval, Matcheri Keshavan and Jukka-Pekka Onnela. 2018. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. *Neuropsychopharmacology*. 1.

[15] Drue H Barrett, Caroline Carney Doebbeling, David A Schwartz, Margaret D Voelker, Kenneth H Falter, Robert F Woolson and Bradley N Doebbeling. 2002. Posttraumatic stress

disorder and self-reported physical health status among US Military personnel serving during the Gulf War period: a population-based study. *Psychosomatics*, 43 (3). 195-205.

[16] Keith C Barton. 2015. Elicitation techniques: Getting people to talk about ideas they don't usually talk about. *Theory & Research in Social Education*, 43 (2). 179-205.

[17] John A Batsis, John A Naslund, Lydia E Gill, Rebecca K Masutani, Nayan Agarwal and Stephen J Bartels. 2016. Use of a wearable activity device in rural older obese adults: A pilot study. *Gerontology and geriatric medicine*, 2. 2333721416678076.

[18] Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone and Jeff Wilson. 2013. Health Mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20 (5). 30.

[19] Linda Birt, Suzanne Scott, Debbie Cavers, Christine Campbell and Fiona Walter. 2016. Member checking: a tool to enhance trustworthiness or merely a nod to validation? *Qualitative health research*, 26 (13). 1802-1811.

[20] Christy A Blevins, Frank W Weathers, Margaret T Davis, Tracy K Witte and Jessica L Domino. 2015. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress*, 28 (6). 489-498.

[21] Michelle J Bovin, Brian P Marx, Frank W Weathers, Matthew W Gallagher, Paola Rodriguez, Paula P Schnurr and Terence M Keane. 2016. Psychometric properties of the PTSD Checklist for Diagnostic and Statistical Manual of Mental Disorders–Fifth Edition (PCL-5) in veterans. *Psychological Assessment*, 28 (11). 1379.

[22] Felicia M Bowens, Patricia A Frye and Warren A Jones. 2010. Health information technology: integration of clinical workflow into meaningful use of electronic health records. *Perspectives in Health Information Management/AHIMA, American Health Information Management Association*, 7 (Fall).

[23] Virginia Braun and Victoria Clarke. 2013. *Successful qualitative research: A practical guide for beginners*. sage.

[24] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, 3 (2). 77-101.

[25] Judith E Brown, Maureen A Murtaugh, David R Jacobs Jr and Helen C Margellos. 2002. Variation in newborn size according to pregnancy weight change by trimester. *The American journal of clinical nutrition*, 76 (1). 205-209.

[26] Ayşe G Büyüktür, Mark S Ackerman, Mark W Newman and Pei-Yao Hung. 2017. Design considerations for semi-automated tracking: self-care plans in spinal cord injury. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 183-192.

[27] Elizabeth Byrnes and James H Johnson. 1981. Change technology and the implementation of automation in mental health care settings. *Behavior Research Methods & Instrumentation*, 13 (4). 573-580.

[28] John S Cacciola, Arthur I Alterman, Dominick DePhilippis, Michelle L Drapkin, Charles Valadez Jr, Natalie C Fala, David Oslin and James R McKay. 2013. Development and initial

evaluation of the Brief Addiction Monitor (BAM). *Journal of Substance Abuse Treatment*, 44 (3). 256-263.

[29] Kelly Caine. 2016. Local standards for sample size at CHI. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, 981-992.

[30] Clara Caldeira, Yu Chen, Lesley Chan, Vivian Pham, Yunan Chen and Kai Zheng. 2017. Mobile apps for mood tracking: an analysis of features and user reviews. In *AMIA Annual Symposium Proceedings*, American Medical Informatics Association, 495.

[31] Juan Luis Carús Candás, Víctor Peláez, Gloria López, Miguel Ángel Fernández, Eduardo Álvarez and Gabriel Díaz. 2014. An automatic data mining method to detect abnormal human behaviour using physical activity measurements. *Pervasive and Mobile Computing*, 15. 228-241.

[32] Domin Chan, Allen D Cheadle, Gayle Reiber, Jürgen Unützer and Edmund F Chaney. 2009. Health care utilization and its costs for depressed veterans with and without comorbid PTSD symptoms. *Psychiatric Services*, 60 (12). 1612-1617.

[33] Eun Kyoung Choe, Saeed Abdullah, Mashfiqui Rabbi, Edison Thomaz, Daniel A Epstein, Felicia Cordeiro, Matthew Kay, Gregory D Abowd, Tanzeem Choudhury and James Fogarty. 2017. Semi-automated tracking: a balanced approach for self-monitoring applications. *IEEE Pervasive Computing*, 16 (1). 74-84.

[34] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche and Dominikus Baur. 2017. Understanding self-reflection: how people reflect on personal data through visual data exploration. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 173-182.

[35] Eun Kyoung Choe, Nicole B Lee, Bongshin Lee, Wanda Pratt and Julie A Kientz. 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM, 1143-1152.

[36] Eun Kyoung Choe, Nicole B Lee, Bongshin Lee, Wanda Pratt and Julie A Kientz. 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1143-1152.

[37] Jenny Chum, Min Suk Kim, Laura Zielinski, Meha Bhatt, Douglas Chung, Sharon Yeung, Kathryn Litke, Kathleen McCabe, Jeff Whattam and Laura Garrick. 2017. Acceptability of the Fitbit in behavioural activation therapy for depression: a qualitative study. *Evidence-based mental health*. ebmental-2017-102763.

[38] Chia-Fang Chung, Jonathan Cook, Elizabeth Bales, Jasmine Zia and Sean A Munson. 2015. More than telemonitoring: health provider use and nonuse of life-log data in irritable bowel syndrome and weight management. *Journal of medical Internet research*, 17 (8).

[39] Chia-Fang Chung, Kristin Dew, Allison Cole, Jasmine Zia, James Fogarty, Julie A Kientz and Sean A Munson. 2016. Boundary negotiating artifacts in personal informatics: Patient-provider collaboration with patient-generated data. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, ACM, 770-786.

[40] Pamela Y Collins, Vikram Patel, Sarah S Joestl, Dana March, Thomas R Insel, Abdallah S Daar, Isabel A Bordin, E Jane Costello, Maureen Durkin and Christopher Fairburn. 2011. Grand challenges in global mental health. *Nature*, 475 (7354). 27.
[41] Lucas Colusso, Cynthia L Bennett, Gary Hsieh and Sean A Munson. 2017. Translational resources: Reducing the gap between academic research and HCI practice. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, ACM, 957-968.
[42] Jesse D Cook, Michael L Prairie and David T Plante. 2017. Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: A comparison against polysomnography and wrist-worn actigraphy. *Journal of affective disorders*, 217. 299-305.
[43] Felicia Cordeiro, Daniel A Epstein, Edison Thomaz, Elizabeth Bales, Arvind K Jagannathan, Gregory D Abowd and James Fogarty. 2015. Barriers and negative nudges: Exploring challenges in food journaling. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 1159-1162.
[44] Kathrin M Cresswell, Ann Blandford and Aziz Sheikh. 2017. Drawing on human factors engineering to evaluate the effectiveness of health information technology. *Journal of the Royal Society of Medicine*, 110 (8). 309-315.
[45] Nathan Crilly, Alan F Blackwell and P John Clarkson. 2006. Graphic elicitation: using research diagrams as interview stimuli. *Qualitative research*, 6 (3). 341-366.
[46] Laura J Damschroder, David C Aron, Rosalind E Keith, Susan R Kirsh, Jeffery A Alexander and Julie C Lowery. 2009. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation science*, 4 (1). 50.
[47] Lorraine Daston. 2007. Objectivity.
[48] Jean-Louis Denis, Yann Hébert, Ann Langley, Daniel Lozeau and Louise-Hélène Trottier. 2002. Explaining diffusion patterns for complex health care innovations. *Health care management review*, 27 (3). 60-73.
[49] Laura Devendorf, Kristina Andersen and Aisling Kelliher. 2020. The fundamental uncertainties of mothering: Finding ways to honor endurance, struggle, and contradiction. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27 (4). 1-24.
[50] Tara Donker, Katherine Petrie, Judy Proudfoot, Janine Clarke, Mary-Rose Birch and Helen Christensen. 2013. Smartphones for smarter delivery of mental health programs: a systematic review. *Journal of medical Internet research*, 15 (11). e247.
[51] Paul Dourish. 2004. What we talk about when we talk about context. *Personal and ubiquitous computing*, 8 (1). 19-30.
[52] Benjamin G Druss, Irving Hwang, Masha Petukhova, Nancy A Sampson, Philip S Wang and Ronald C Kessler. 2009. Impairment in role functioning in mental and chronic medical disorders in the United States: results from the National Comorbidity Survey Replication. *Molecular psychiatry*, 14 (7). 728.
[53] Benjamin G Druss, Steven C Marcus, Robert A Rosenheck, Mark Olfson, Terri Tanielian and Harold A Pincus. 2000. Understanding disability in mental and general medical conditions. *American Journal of Psychiatry*, 157 (9). 1485-1491.

[54] Elizabeth V Eikey and Madhu C Reddy. 2017. " It's Definitely Been a Journey" A Qualitative Study on How Women with Eating Disorders Use Weight Loss Apps. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, 642-654.
[55] Elizabeth Victoria Eikey, Clara Marques Caldeira, Mayara Costa Figueiredo, Yunan Chen, Jessica L Borelli, Melissa Mazmanian and Kai Zheng. 2021. Beyond self-reflection: introducing the concept of rumination in personal informatics. *Personal and Ubiquitous Computing*. 1-16.
[56] Chris Elsden, David S Kirk and Abigail C Durrant. 2016. A quantified past: Toward design for remembering with personal informatics. *Human–Computer Interaction*, 31 (6). 518-557.
[57] Daniel A Epstein, Alan Borning and James Fogarty. 2013. Fine-grained sharing of sensed physical activity: a value sensitive approach. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, 489-498.
[58] Daniel A Epstein, Nicole B Lee, Jennifer H Kang, Elena Agapie, Jessica Schroeder, Laura R Pina, James Fogarty, Julie A Kientz and Sean Munson. 2017. Examining menstrual tracking to inform the design of personal informatics tools. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 6876-6888.
[59] Daniel A Epstein, An Ping, James Fogarty and Sean A Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 731-742.
[60] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty and Sean Munson. 2014. Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In *Proceedings of the 2014 conference on Designing interactive systems*, 667-676.
[61] Iris Epstein, Bonnie Stevens, Patricia McKeever and Sylvain Baruchel. 2006. Photo elicitation interview (PEI): Using photos to elicit children's perspectives. *International journal of qualitative methods*, 5 (3). 1-11.
[62] Maria Faurholt-Jepsen, Jonas Busk, Mads Frost, Maj Vinberg, Ellen M Christensen, Ole Winther, Jakob Eyvind Bardram and Lars V Kessing. 2016. Voice analysis as an objective state marker in bipolar disorder. *Translational psychiatry*, 6 (7). e856.
[63] Ewan Ferlie, John Gabbay, Louise Fitzgerald, Louise Locock and Sue Dopson. 2001. Evidence-based medicine and organisational change: an overview of some recent qualitative research.
[64] Joseph Firth and John Torous. 2015. Smartphone apps for schizophrenia: a systematic review. *JMIR mHealth and uHealth*, 3 (4). e102.
[65] Joseph Firth, John Torous, Jennifer Nicholas, Rebekah Carney, Abhishek Pratap, Simon Rosenbaum and Jerome Sarris. 2017. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry*, 16 (3). 287-298.
[66] Geraldine Fitzpatrick and Gunnar Ellingsen. 2013. A review of 25 years of CSCW research in healthcare: contributions, challenges and future agendas. *Computer Supported Cooperative Work (CSCW)*, 22 (4-6). 609-665.
[67] Edna B Foa, Anke Ehlers, David M Clark, David F Tolin and Susan M Orsillo. 1999. The posttraumatic cognitions inventory (PTCI): Development and validation. *Psychological assessment*, 11 (3). 303.

[68] APA Task Force. 2006. APA presidential task force on evidence based practice. *American Psychologist*, 61. 271-285.

[69] John C Fortney, Jürgen Unützer, Glenda Wrenn, Jeffrey M Pyne, G Richard Smith, Michael Schoenbaum and Henry T Harbin. 2016. A tipping point for measurement-based care. *Psychiatric Services*, 68 (2). 179-188.

[70] Sarah E Fox, Amanda Menking, Jordan Eschler and Uba Backonja. 2020. Multiples Over Models: Interrogating the Past and Collectively Reimagining the Future of Menstrual Sensemaking. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27 (4). 1-24.

[71] Jerome D Frank and Julia B Frank. 1993. *Persuasion and healing: A comparative study of psychotherapy*. JHU Press.

[72] Susan M Frayne, Victor Y Chiu, Samina Iqbal, Eric A Berg, Kaajal J Laungani, Ruth C Cronkite, Joanne Pavao and Rachel Kimerling. 2011. Medical care needs of returning veterans with PTSD: their other burden. *Journal of general internal medicine*, 26 (1). 33-39.

[73] Bridget Freisthler, Sharon Lipperman-Kreda, Melina Bersamin and Paul J Gruenewald. 2014. Tracking the when, where, and with whom of alcohol use: Integrating ecological momentary assessment and geospatial data to examine risk for alcohol-related problems. *Alcohol Research: Current Reviews*, 36 (1). 29.

[74] TW Frick and CM Reigeluth. 1999. Formative research: A methodology for creating and improving design theories. *Instructional-design theories and models: A new paradigm of instructional theory*, 2. 633-652.

[75] Matthew J Friedman, Terence M Keane and Patricia A Resick. 2007. *Handbook of PTSD: Science and practice*. Guilford Press.

[76] Jessica J Fulton, Patrick S Calhoun, H Ryan Wagner, Amie R Schry, Lauren P Hair, Nicole Feeling, Eric Elbogen and Jean C Beckham. 2015. The prevalence of posttraumatic stress disorder in Operation Enduring Freedom/Operation Iraqi Freedom (OEF/OIF) Veterans: a meta-analysis. *Journal of anxiety disorders*, 31. 98-107.

[77] David Gauntlett. 2007. *Creative explorations: New approaches to identities and audiences*. Routledge.

[78] Asma Ghandeharioun, Szymon Fedor, Lisa Sangermano, Dawn Ionescu, Jonathan Alpert, Chelsea Dale, David Sontag and Rosalind Picard. 2017. Objective assessment of depressive symptoms with machine learning and wearable sensors data. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, IEEE, 325-332.

[79] Paul Gill, Kate Stewart, Elizabeth Treasure and Barbara Chadwick. 2008. Methods of data collection in qualitative research: interviews and focus groups. *British dental journal*, 204 (6). 291-295.

[80] Russell E Glasgow, Cynthia Vinson, David Chambers, Muin J Khoury, Robert M Kaplan and Christine Hunter. 2012. National Institutes of Health approaches to dissemination and implementation science: current and future directions. *American journal of public health*, 102 (7). 1274-1281.

[81] Erving Goffman. 1978. *The presentation of self in everyday life*. Harmondsworth London.

[82] Connie Golsteijn, Elise Van Den Hoven, David Frohlich and Abigail Sellen. 2012. Towards a more cherishable digital object. In *Proceedings of the designing interactive systems conference*, 655-664.
[83] Rúben Gouveia, Evangelos Karapanos and Marc Hassenzahl. 2015. How do we engage with activity trackers? A longitudinal study of Habito. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 1305-1316.
[84] Esther Granado-Font, Gemma Flores-Mateo, Mar Sorlí-Aguilar, Xavier Montaña-Carreras, Carme Ferre-Grau, Maria-Luisa Barrera-Uriarte, Eulàlia Oriol-Colominas, Cristina Rey-Reñones, Iolanda Caules and Eva-María Satué-Gracia. 2015. Effectiveness of a Smartphone application and wearable device for weight loss in overweight or obese primary care patients: protocol for a randomised controlled trial. *BMC Public Health*, 15 (1). 531.
[85] Saul Greenberg and Bill Buxton. 2008. Usability evaluation considered harmful (some of the time). In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 111-120.
[86] Trisha Greenhalgh, Glenn Robert, Fraser Macfarlane, Paul Bate and Olivia Kyriakidou. 2004. Diffusion of innovations in service organizations: systematic review and recommendations. *The Milbank Quarterly*, 82 (4). 581-629.
[87] Carla F Griggio, Midas Nouwens, Joanna Mcgrenere and Wendy E Mackay. 2019. Augmenting couples' communication with lifelines: Shared timelines of mixed contextual information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1-13.
[88] Rebecca Gulotta, William Odom, Jodi Forlizzi and Haakon Faste. 2013. Digital artifacts as legacy: exploring the lifespan and value of digital data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1813-1822.
[89] Oliver L Haimson, Bryan Semaan, Brianna Dym, Joey Chiao-Yin Hsiao, Daniel Herron and Wendy Moncur. 2019. Life Transitions and Social Technologies: Research and Design for Times of Life Change. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, 480-486.
[90] Shefali Haldar, Sonali R Mishra, Maher Khelifi, Ari H Pollack and Wanda Pratt. 2019. The Value of Patient-Peer Support in Improving Hospital Safety. In *ITCH*, 152-156.
[91] Kelli Jane K Harding, A John Rush, Melissa Arbuckle, Madhukar H Trivedi and Harold Alan Pincus. 2011. Measurement-based care in psychiatric practice: a policy framework for implementation. *The Journal of clinical psychiatry*.
[92] Richard Harper, David Randall, Nicola Smyth, Carwyn Evans, Lisa Heledd and R Moore. 2008. The past is a different place: they do things differently there. In *Proceedings of the 7th ACM conference on Designing interactive systems*, 271-280.
[93] Michael I Harrison, Ross Koppel and Shirly Bar-Lev. 2007. Unintended consequences of information technologies in health care—an interactive sociotechnical analysis. *Journal of the American medical informatics Association*, 14 (5). 542-549.
[94] Margaret M Harvey, Sheila AM Rauch, Alyson K Zalta, Jo Sornborger, Mark H Pollack, Barbara O Rothbaum, Lauren M Laifer and Naomi M Simon. 2017. Intensive Treatment Models

to Address Posttraumatic Stress Among Post-9/11 Warriors: The Warrior Care Network. *Focus*, 15 (4). 378-383.

[95] National Institute of Health.

[96] Christian Heath, Jon Hindmarsh and Paul Luff. 2010. *Video in qualitative research*. Sage Publications.

[97] Christine Hedman, Tarja Pohjasvaara, Uolevi Tolonen, AS Suhonen-Malm and VV Myllylä. 2002. Effects of pregnancy on mothers' sleep. *Sleep medicine*, 3 (1). 37-42.

[98] Karen L Henwood and Nick F Pidgeon. 1992. Qualitative research and psychological theorizing. *British journal of psychology*, 83 (1). 97-111.

[99] Daniel Herron, Nazanin Andalibi, Oliver Haimson, Wendy Moncur and Elise van den Hoven. 2016. HCI and sensitive life experiences. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, 1-3.

[100] Geoffrey M Hodgson. 2008. The concept of a routine. *Handbook of organizational routines*, 15.

[101] Geoffrey M Hodgson. 1997. The ubiquity of habits and rules. *Cambridge journal of economics*, 21 (6). 663-684.

[102] Trevor Hogan, Uta Hinrichs and Eva Hornecker. 2015. The elicitation interview technique: capturing people's experiences of data representations. *IEEE transactions on visualization and computer graphics*, 22 (12). 2579-2593.

[103] Immy Holloway and Kathleen Galvin. 2016. *Qualitative research in nursing and healthcare*. John Wiley & Sons.

[104] Matthew K Hong, Lauren Wilcox, Daniel Machado, Thomas A Olson and Stephen F Simoneaux. 2016. Care partnerships: Toward technology to support teens' participation in their health care. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5337-5349.

[105] Noura Howell, Laura Devendorf, Tomás Alfonso Vega Gálvez, Rundong Tian and Kimiko Ryokai. 2018. Tensions of data-driven reflection: A case study of real-time emotional biosensing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, 431.

[106] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang and Robert Woodbury. 2014. Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 21 (3). 420-433.

[107] Rhidian Hughes. 1998. Considering the vignette technique and its application to a study of drug injecting and HIV risk and safer behaviour. *Sociology of Health & Illness*, 20 (3). 381-400.

[108] Natalie E Hundt, Juliette M Mott, Shannon R Miles, Jennifer Arney, Jeffrey A Cully and Melinda A Stanley. 2015. Veterans' perspectives on initiating evidence-based psychotherapy for posttraumatic stress disorder. *Psychological Trauma: Theory, Research, Practice, and Policy*, 7 (6). 539.

[109] IDC. 2019. IDC Reports Strong Growth in the Worldwide Wearables Market, Led by Holiday Shipments of Smartwatches, Wrist Bands, and Ear-Worn Devices.

[110] S. S. Intille. 2004. Ubiquitous computing technology for just-in-time motivation of behavior change. *Stud Health Technol Inform*, 107 (Pt 2). 1434-1437.
[111] Maia L Jacobs, James Clawson and Elizabeth D Mynatt. 2017. Articulating a Patient-Centered Design Space for Cancer Journeys. *EAI endorsed transactions on pervasive health and technology*, 3 (9).
[112] Maia L Jacobs, James Clawson and Elizabeth D Mynatt. 2015. Comparing health information sharing preferences of cancer patients, doctors, and navigators. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 808-818.
[113] Maia L Jacobs, James Clawson and Elizabeth D Mynatt. 2014. 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*, 663-672.
[114] Luis G Jaimes, Martin Llofriu and Andrew Raij. 2014. 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*, 197-203.
[115] Aaron Jarden. 2017. An Interview with Victor J. Strecher. *International Journal of Wellbeing*, 7 (2).
[116] Roger Jones. 1995. Why do qualitative research?, British Medical Journal Publishing Group.
[117] Heidi Julien, Lisa M Given and Anna Opryshko. 2013. Photovoice: A promising method for studies of individuals' information practices. *Library & Information Science Research*, 35 (4). 257-263.
[118] Mahdokht Kalantari. 2017. Consumers' adoption of wearable technologies: literature review, synthesis, and future research agenda. *International Journal of Technology Marketing*, 12 (3). 274-307.
[119] Hanna Kallio, Anna-Maija Pietilä, Martin Johnson and Mari Kangasniemi. 2016. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of advanced nursing*, 72 (12). 2954-2965.
[120] Ravi Karkar, James Fogarty, Julie A Kientz, Sean A Munson, Roger Vilardaga and Jasmine Zia. 2015. Opportunities and challenges for self-experimentation in self-tracking. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, 991-996.
[121] Ravi Karkar, Jasmine Zia, Roger Vilardaga, Sonali R Mishra, James Fogarty, Sean A Munson and Julie A Kientz. 2016. A framework for self-experimentation in personalized health. *Journal of the American Medical Informatics Association*, 23 (3). 440-448.
[122] Anne Kaun and Fredrik Stiernstedt. 2014. Facebook time: Technological and institutional affordances for media memories. *New Media & Society*, 16 (7). 1154-1168.
[123] Shannon M Kehle, Nancy Greer, Indulis Rutks and Timothy Wilt. 2011. Interventions to improve veterans' access to care: a systematic review of the literature. *Journal of general internal medicine*, 26 (2). 689.

[124] Shannon M Kehle-Forbes, Laura A Meis, Michele R Spoont and Melissa A Polusny. 2016. Treatment initiation and dropout from prolonged exposure and cognitive processing therapy in a VA outpatient clinic. *Psychological Trauma: Theory, Research, Practice, and Policy*, 8 (1). 107.
[125] Christina Kelley, Bongshin Lee and Lauren Wilcox. 2017. Self-tracking for mental wellness: understanding expert perspectives and student experiences. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 629-641.
[126] Susan E Kelly, I Bourgeault and R Dingwall. 2010. Qualitative interviewing techniques and styles. *The SAGE handbook of qualitative methods in health research*. 307-326.
[127] Edmund Keogh and Gordon JG Asmundson. 2004. Negative affectivity, catastrophizing, and anxiety sensitivity. *Understanding and treating fear of pain*. 91-115.
[128] Ronald C Kessler. 2000. Posttraumatic stress disorder: the burden to the individual and to society. *The Journal of clinical psychiatry*.
[129] Ronald C Kessler, Patricia Berglund, Olga Demler, Robert Jin, Kathleen R Merikangas and Ellen E Walters. 2005. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of general psychiatry*, 62 (6). 593-602.
[130] Ronald C Kessler, Wai Tat Chiu, Olga Demler and Ellen E Walters. 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of general psychiatry*, 62 (6). 617-627.
[131] Amy M Kilbourne, Mary S Neumann, Harold A Pincus, Mark S Bauer and Ronald Stall. 2007. Implementing evidence-based interventions in health care: application of the replicating effective programs framework. *Implementation Science*, 2 (1). 42.
[132] Young-Ho Kim, Jae Ho Jeon, Bongshin Lee, Eun Kyoung Choe and Jinwook Seo. 2017. OmniTrack: A flexible self-tracking approach leveraging semi-automated tracking. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1 (3). 1-28.
[133] Bran Knowles, Alison Smith-Renner, Forough Poursabzi-Sangdeh, Di Lu and Halimat Alabi. 2018. Uncertainty in current and future health wearables. *Communications of the ACM*, 61 (12). 62-67.
[134] Judy Kopp. 1988. Self-monitoring: A literature review of research and practice. In *Social Work Research and Abstracts*, Oxford University Press, 8-20.
[135] Ross Koppel and David Kreda. 2009. Health care information technology vendors'"hold harmless" clause: implications for patients and clinicians. *JAMA*, 301 (12). 1276-1278.
[136] Kurt Kroenke and Robert L Spitzer. 2002. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric annals*, 32 (9). 509-515.
[137] Edward S Kubany, Stephen N Haynes, Francis R Abueg, Frederic P Manke, Jerry M Brennan and Catherine Stahura. 1996. Development and validation of the Trauma-Related Guilt Inventory (TRGI). *Psychological Assessment*, 8 (4). 428.
[138] Amanda Lazar, Christian Koehler, Joshua Tanenbaum and David H Nguyen. 2015. Why we use and abandon smart devices. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 635-646.
[139] Allen S Lee and Richard L Baskerville. 2003. Generalizing generalizability in information systems research. *Information systems research*, 14 (3). 221-243.

[140] Cara C Lewis, Meredith Boyd, Ajeng Puspitasari, Elena Navarro, Jacqueline Howard, Hannah Kassab, Mira Hoffman, Kelli Scott, Aaron Lyon and Susan Douglas. 2019. Implementing measurement-based care in behavioral health: A review. *JAMA psychiatry*, 76 (3). 324-335.

[141] Ian Li, Anind Dey and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 557-566.

[142] Ian Li, Anind K Dey and Jodi Forlizzi. 2011. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing*, ACM, 405-414.

[143] Ian Li, Anind K Dey and Jodi Forlizzi. 2012. Using context to reveal factors that affect physical activity. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 19 (1). 1-21.

[144] Bob Kummerfeld Lie Ming Tang, Judy Kay. 2018. Decades of data: will it be useless?

[145] Giuseppe Lippi, Brandon M Henry and Fabian Sanchis-Gomar. 2020. Physical inactivity and cardiovascular disease at the time of coronavirus disease 2019 (COVID-19). *European Journal of Preventive Cardiology*. 2047487320916823.

[146] Ross J Lordon, Sean P Mikles, Laura Kneale, Heather L Evans, Sean A Munson, Uba Backonja and William B Lober. 2020. How patient-generated health data and patient-reported outcomes affect patient–clinician relationships: A systematic review. *Health informatics journal*. 1460458220928184.

[147] Deborah Lupton. 2020. CARING DATAVEILLANCE. *The Routledge Companion to Digital Media and Children*.

[148] Deborah Lupton. 2020. Caring dataveillance: Women's use of apps to monitor pregnancy and children. in *The Routledge companion to digital media and children*, Routledge, 393-402.

[149] Deborah Lupton. 2020. 'Not the Real Me': Social Imaginaries of Personal Data Profiling. *Cultural Sociology*. 1749975520939779.

[150] Deborah Lupton. 2013. Quantifying the body: monitoring and measuring health in the age of mHealth technologies. *Critical Public Health*, 23 (4). 393-403.

[151] Deborah Lupton. 2014. Self-tracking cultures: towards a sociology of personal informatics. In *Proceedings of the 26th Australian computer-human interaction conference on designing futures: The future of design*, 77-86.

[152] Deborah Lupton. 2017. Self-tracking, health and medicine, Taylor & Francis.

[153] Deborah Lupton. 2016. You are your data: Self-tracking practices and concepts of data. in *Lifelogging*, Springer, 61-79.

[154] Yuliya Lutchyn, Paul Johns, Asta Roseway and Mary Czerwinski. 2015. MoodTracker: Monitoring collective emotions in the workplace. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, IEEE, 295-301.

[155] Haley MacLeod, Anthony Tang and Sheelagh Carpendale. 2013. Personal informatics in chronic illness management. in *Proceedings of Graphics Interface 2013*, Citeseer, 149-156.

[156] Lena Mamykina, Elizabeth M Heitkemper, Arlene M Smaldone, Rita Kukafka, Heather J Cole-Lewis, Patricia G Davidson, Elizabeth D Mynatt, Andrea Cassells, Jonathan N Tobin and

George Hripcsak. 2017. Personal discovery in diabetes self-management: discovering cause and effect using self-monitoring data. *Journal of biomedical informatics*, 76. 1-8.
[157] Lena Mamykina, Elizabeth D Mynatt and David R Kaufman. 2006. Investigating health management practices of individuals with diabetes. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, ACM, 927-936.
[158] Lena Mamykina, Elizabeth Mynatt, Patricia Davidson and Daniel Greenblatt. 2008. MAHI: investigation of social scaffolding for reflective thinking in diabetes management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 477-486.
[159] Michael Massimi, Jill P Dimond and Christopher A Le Dantec. 2012. Finding a new normal: the role of technology in life disruptions. In *Proceedings of the acm 2012 conference on computer supported cooperative work*, 719-728.
[160] Mark Matthews, Elizabeth Murnane and Jaime Snyder. 2017. Quantifying the Changeable Self: The role of self-tracking in coming to terms with and managing bipolar disorder. *Human–Computer Interaction*, 32 (5-6). 413-446.
[161] Elina Mattila, Juha Pärkkä, Marion Hermersdorf, Jussi Kaasinen, Janne Vainio, Kai Samposalo, Juho Merilahti, Juha Kolari, Minna Kulju and Raimo Lappalainen. 2008. Mobile diary for wellness management—results on usage and usability in two user studies. *IEEE Transactions on information technology in biomedicine*, 12 (4). 501-512.
[162] Oscar Mayora, Bert Arnrich, Jakob Bardram, Carsten Dräger, Andrea Finke, Mads Frost, Silvia Giordano, Franz Gravenhorst, Agnes Grunerbl and Christian Haring. 2013. Personal health systems for bipolar disorder anecdotes, challenges and lessons learnt from monarca project. In *2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*, IEEE, 424-429.
[163] Jingbo Meng, Syed Ali Hussain, David C Mohr, Mary Czerwinski and Mi Zhang. 2018. Exploring User Needs for a Mobile Behavioral-Sensing Technology for Depression Management: Qualitative Study. *Journal of medical Internet research*, 20 (7).
[164] Helena M Mentis, Anita Komlodi, Katrina Schrader, Michael Phipps, Ann Gruber-Baldini, Karen Yarbrough and Lisa Shulman. 2017. Crafting a view of self-tracking data in the clinical visit. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ACM, 5800-5812.
[165] Helena M Mentis, Rita Shewbridge, Sharon Powell, Melissa Armstrong, Paul Fishman and Lisa Shulman. 2016. Co-interpreting movement with sensors: assessing Parkinson’s patients’ deep brain stimulation programming. *Human–Computer Interaction*, 31 (3-4). 227-260.
[166] Jochen Meyer, Judy Kay, Daniel A Epstein, Parisa Eslambolchilar and Lie Ming Tang. 2020. A life of data: Characteristics and challenges of very long term self-tracking for health and wellness. *ACM Transactions on Computing for Healthcare*, 1 (2). 1-4.
[167] Baukje Miedema, Ryan Hamilton and Julie Easley. 2007. From “invincibility” to “normalcy”: Coping strategies of young adults during the cancer journey. *Palliative & Supportive Care*, 5 (1). 41-49.
[168] Matthew B Miles and A Michael Huberman. 1984. Drawing valid meaning from qualitative data: Toward a shared craft. *Educational researcher*, 13 (5). 20-30.

[169] Sonali R Mishra, Predrag Klasnja, John MacDuffie Woodburn, Eric B Hekler, Larsson Omberg, Michael Kellen and Lara Mangravite. 2019. Supporting coping with parkinson's disease through self tracking. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1-16.

[170] Sonali R Mishra, Woody MacDuffie, Predrag Klasnja, Eric Hekler, Larsson Omberg, Michael Kellen and Lara Mangravite. Attitudes Towards Long-Term Tracking in Parkinson's Disease. *Ann Arbor*, 1001. 48109.

[171] Sonali R Mishra, Andrew D Miller, Shefali Haldar, Maher Khelifi, Jordan Eschler, Rashmi G Elera, Ari H Pollack and Wanda Pratt. 2018. Supporting collaborative health tracking in the hospital: patients' perspectives. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1-14.

[172] Sonali R Mishra, Andrew D Miller, Shefali Haldar, Maher Khelifi, Jordan Eschler, Rashmi G Elera, Ari H Pollack and Wanda Pratt. 2018. Supporting Collaborative Health Tracking in the Hospital: Patients' Perspectives. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, 650.

[173] David C Mohr, Mi Zhang and Stephen M Schueller. 2017. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. *Annual review of clinical psychology*, 13. 23-47.

[174] Timothy K Monk, Joseph F Flaherty, Ellen Frank, Kathleen Hoskinson and David J Kupfer. 1990. The Social Rhythm Metric: An instrument to quantify the daily rhythms of life. *Journal of Nervous and Mental Disease*.

[175] Raeanne C Moore, Colin A Depp, Julie Loebach Wetherell and Eric J Lenze. 2016. Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. *Journal of psychiatric research*, 75. 116-123.

[176] Bjorn E Munkvold. 1999. Challenges of IT implementation for supporting collaboration in distributed organizations. *European Journal of Information Systems*, 8 (4). 260-272.

[177] Scott A Murray, Marilyn Kendall, Kirsty Boyd and Aziz Sheikh. 2005. Illness trajectories and palliative care. *Bmj*, 330 (7498). 1007-1011.

[178] Bonnie A Nardi. 1996. Studying context: A comparison of activity theory, situated action models, and distributed cognition. *Context and consciousness: Activity theory and human-computer interaction*, 69102.

[179] Ada Ng, Madhu Reddy, Alyson K Zalta and Stephen M Schueller. 2018. Veterans' Perspectives on Fitbit Use in Treatment for Post-Traumatic Stress Disorder: An Interview Study. *JMIR mental health*, 5 (2). e10415.

[180] Francisco Nunes and Geraldine Fitzpatrick. 2018. Understanding the mundane nature of self-care: Ethnographic accounts of people living with Parkinson's. In *proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1-15.

[181] Francisco Nunes, Nervo Verdezoto, Geraldine Fitzpatrick, Morten Kyng, Erik Grönvall and Cristiano Storni. 2015. Self-care technologies in HCI: Trends, tensions, and opportunities. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 22 (6). 1-45.

[182] Aisling Ann O'Kane, Yvonne Rogers and Ann E Blandford. 2014. Gaining empathy for non-routine mobile device use through autoethnography. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, 987-990.

[183] Wanda J Orlikowski. 1993. Learning from notes: Organizational issues in groupware implementation.

[184] Wanda J Orlikowski and JoAnne Yates. 2002. It's about time: Temporal structuring in organizations. *Organization science*, 13 (6). 684-700.

[185] Aswati Panicker, Kavya Basu and Chia-Fang Chung. 2020. Changing Roles and Contexts: Symbolic Interactionism in the Sharing of Food and Eating Practices between Remote, Intergenerational Family Members. *Proceedings of the ACM on Human-Computer Interaction*, 4 (CSCW1). 1-19.

[186] Mika Pantzar and Minna Ruckenstein. 2017. Living the metrics: Self-tracking and situated objectivity. *Digital health*, 3. 2055207617712590.

[187] Dana Pavel, Vic Callaghan, Francisco Sepulveda, Michael Gardner and Anind K Dey. 2012. The story of our lives: From sensors to stories in self-monitoring systems. In *2012 4th Computer Science and Electronic Engineering Conference (CEEC)*, IEEE, 71-76.

[188] Veljko Pejovic and Mirco Musolesi. 2015. Anticipatory mobile computing: A survey of the state of the art and research challenges. *ACM Computing Surveys (CSUR)*, 47 (3). 1-29.

[189] Justin Petelka, Lucy Van Kleunen, Liam Albright, Elizabeth Murnane, Stephen Voida and Jaime Snyder. 2020. Being (In) Visible: Privacy, Transparency, and Disclosure in the Self-Management of Bipolar Disorder. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-14.

[190] Wanda Pratt, Madhu C Reddy, David W McDonald, Peter Tarczy-Hornoch and John H Gennari. 2004. Incorporating ideas from computer-supported cooperative work. *Journal of biomedical informatics*, 37 (2). 128-137.

[191] Mashfiqui Rabbi, Shahid Ali, Tanzeem Choudhury and Ethan Berke. 2011. Passive and in-situ assessment of mental and physical well-being using mobile sensors. In *Proceedings of the 13th international conference on Ubiquitous computing*, 385-394.

[192] Mashfiqui Rabbi, Min Hane Aung, Mi Zhang and Tanzeem Choudhury. 2015. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 707-718.

[193] Shriti Raj, Kelsey Toporski, Ashley Garrity, Joyce M Lee and Mark W Newman. 2019. My blood sugar is higher on the weekends: Finding a Role for Context and Context-Awareness in the Design of Health Self-Management Technology. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, 119.

[194] Ruth Ravichandran, Sang-Wha Sien, Shwetak N Patel, Julie A Kientz and Laura R Pina. 2017. Making sense of sleep sensors: How sleep sensing technologies support and undermine sleep health. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ACM, 6864-6875.

[195] Christian Remy, Oliver Bates, Alan Dix, Vanessa Thomas, Mike Hazas, Adrian Friday and Elaine M Huang. 2018. Evaluation beyond usability: Validating sustainable HCI research. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1-14.

[196] Donghao Ren, Matthew Brehmer, Bongshin Lee, Tobias Höllerer and Eun Kyoung Choe. 2017. Chartaccent: Annotation for data-driven storytelling. In *2017 IEEE Pacific Visualization Symposium (PacificVis)*, Ieee, 230-239.

[197] Patricia A Resick, Candice M Monson and Kathleen M Chard. 2016. *Cognitive processing therapy for PTSD: A comprehensive manual*. Guilford Publications.

[198] Gary Rolfe. 2006. Validity, trustworthiness and rigour: quality and the idea of qualitative research. *Journal of advanced nursing*, 53 (3). 304-310.

[199] John Rooksby, Alistair Morrison and Dave Murray-Rust. 2019. Student Perspectives on Digital Phenotyping: The Acceptability of Using Smartphone Data to Assess Mental Health. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, 425.

[200] John Rooksby, Mattias Rost, Alistair Morrison and Matthew Chalmers. 2014. Personal tracking as lived informatics. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 1163-1172.

[201] Alan Rozet, Ian M Kronish, Joseph E Schwartz and Karina W Davidson. 2019. Using machine learning to derive just-in-time and personalized predictors of stress: observational study bridging the gap between nomothetic and ideographic approaches. *Journal of medical Internet research*, 21 (4). e12910.

[202] Herbert J Rubin and Irene S Rubin. 2011. *Qualitative interviewing: The art of hearing data*. sage.

[203] Minna Ruckenstein and Natasha Dow Schüll. 2017. The datafication of health. *Annual Review of Anthropology*, 46. 261-278.

[204] Koustuv Saha, Ayse E Bayraktaroglu, Andrew T Campbell, Nitesh V Chawla, Munmun De Choudhury, Sidney K D'Mello, Anind K Dey, Ge Gao, Julie M Gregg and Krithika Jagannath. 2019. Social media as a passive sensor in longitudinal studies of human behavior and wellbeing. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, CS12.

[205] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty and Kristina Höök. 2019. HCI and Affective Health: Taking stock of a decade of studies and charting future research directions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, 245.

[206] Jomara Sandbulte, Jordan Beck, Eun Kyoung Choe and John M Carroll. 2019. Turning Points: Motivating Intergenerational Families to Engage on Sustainable Health Information Sharing. In *International Conference on Information*, Springer, 741-753.

[207] Kelly Santomas. 2020. Passive Monitoring of Physical Activity Using a Fitbit Charge 3 in Individuals at Risk for Depression.

[208] Hillol Sarker, Matthew Tyburski, Md Mahbubur Rahman, Karen Hovsepian, Moushumi Sharmin, David H Epstein, Kenzie L Preston, C Debra Furr-Holden, Adam Milam and Inbal Nahum-Shani. 2016. Finding significant stress episodes in a discontinuous time series of

rapidly varying mobile sensor data. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, 4489-4501.
[209] Corina Sas, Steve Whittaker, Steven Dow, Jodi Forlizzi and John Zimmerman. 2014. Generating implications for design through design research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1971-1980.
[210] Zachary Schmitt and Svetlana Yarosh. 2018. Participatory design of technologies to support recovery from substance use disorders. *Proceedings of the ACM on Human-Computer Interaction*, 2 (CSCW). 1-27.
[211] Elisabeth Schramm, Dietrich van Calker, Petra Dykierek, Klaus Lieb, Sabine Kech D Clin Psych, Ingo Zobel D Clin Psych, Rainer Leonhart D Clin Psych and Mathias Berger. 2007. An intensive treatment program of interpersonal psychotherapy plus pharmacotherapy for depressed inpatients: acute and long-term results. *American Journal of Psychiatry*, 164 (5). 768-777.
[212] Jessica Schroeder, Jane Hoffswell, Chia-Fang Chung, James Fogarty, Sean Munson and Jasmine Zia. 2017. Supporting patient-provider collaboration to identify individual triggers using food and symptom journals. In *CSCW: proceedings of the Conference on Computer-Supported Cooperative Work. Conference on Computer-Supported Cooperative Work*, NIH Public Access, 1726.
[213] Jessica Schroeder, Ravi Karkar, Natalia Murinova, James Fogarty and Sean A Munson. 2019. Examining opportunities for goal-directed self-tracking to support chronic condition management. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3 (4). 1-26.
[214] Karen H Seal, Shira Maguen, Beth Cohen, Kristian S Gima, Thomas J Metzler, Li Ren, Daniel Bertenthal and Charles R Marmar. 2010. VA mental health services utilization in Iraq and Afghanistan veterans in the first year of receiving new mental health diagnoses. *Journal of traumatic stress*, 23 (1). 5-16.
[215] Arash Shaban-Nejad, Martin Michalowski and David L Buckeridge. 2018. Health intelligence: how artificial intelligence transforms population and personalized health, Nature Publishing Group.
[216] Michael Shapiro, Douglas Johnston, Jonathan Wald and Donald Mon. 2012. Patient-generated health data. *RTI International, April*.
[217] Moushumi Sharmin, Andrew Raij, David Epstien, Inbal Nahum-Shani, J Gayle Beck, Sudip Vhaduri, Kenzie Preston and Santosh Kumar. 2015. 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*, 505-516.
[218] Tamar Sharon. 2017. Self-tracking for health and the quantified self: Re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. *Philosophy & Technology*, 30 (1). 93-121.
[219] Yunzhou Shi, Megan Manco, Dominique Moyal, Gil Huppert, Hitoshi Araki, Anthony Banks, Hemant Joshi, Richard McKenzie, Alex Seewald and Guy Griffin. 2018. Soft, stretchable, epidermal sensor with integrated electronics and photochemistry for measuring personal UV exposures. *PloS one*, 13 (1). e0190233.

[220] Saul Shiffman, Arthur A Stone and Michael R Hufford. 2008. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.*, 4. 1-32.

[221] Andra Siibak. 2009. Constructing the self through the photo selection-visual impression management on social networking websites. *Cyberpsychology: Journal of psychosocial research on cyberspace*, 3 (1).

[222] Marjorie Skubic, Rainer Dane Guevara and Marilyn Rantz. 2015. Automated health alerts using in-home sensor data for embedded health assessment. *IEEE journal of translational engineering in health and medicine*, 3. 1-11.

[223] Jaime Snyder, Elizabeth Murnane, Caitie Lustig and Stephen Voida. 2019. Visually Encoding the Lived Experience of Bipolar Disorder. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, 133.

[224] Robert L Spitzer, Kurt Kroenke, Janet BW Williams and Bernd Löwe. 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*, 166 (10). 1092-1097.

[225] Sol Steinmetz. 1999. *Random House Webster's unabridged dictionary*. Random House.

[226] Jina Suh, Spencer Williams, Jesse R Fann, James Fogarty, Amy M Bauer and Gary Hsieh. 2020. Parallel Journeys of Patients with Cancer and Depression: Challenges and Opportunities for Technology-Enabled Collaborative Care. *Proceedings of the ACM on Human-Computer Interaction*, 4 (CSCW1). 1-36.

[227] Shaoxiong Sun, Amos Folarin, Yatharth Ranjan, Zulqarnain Rashid, Pauline Conde, Nicholas Cummins, Faith Matcham, Gloria Dalla Costa, Letizia Leocani and Per Soelberg Sørensen. 2020. Using smartphones and wearable devices to monitor behavioural changes during COVID-19. *arXiv preprint arXiv:2004.14331*.

[228] Melanie Swan. 2009. Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. *International journal of environmental research and public health*, 6 (2). 492-525.

[229] M Clare Taylor. 2005. Interviewing. *Qualitative research in health care*. 39-55.

[230] Claudia Thiele, Anton-Rupert Laireiter and Urs Baumann. 2002. Diaries in clinical psychology and psychotherapy: A selective review. *Clinical Psychology & Psychotherapy*, 9 (1). 1-37.

[231] Anja Thieme, Danielle Belgrave and Gavin Doherty. Machine Learning in Mental Health: A Systematic Review of the HCI Literature to Support Effective ML System Design. *ACM Transactions on Computer-Human Interaction (TOCHI).*

[232] Anja Thieme, Jayne Wallace, Paula Johnson, John McCarthy, Siân Lindley, Peter Wright, Patrick Olivier and Thomas D Meyer. 2013. Design to promote mindfulness practice and sense of self for vulnerable women in secure hospital services. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2647-2656.

[233] Edison Thomaz, Irfan Essa and Gregory D Abowd. 2015. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 1029-1040.

[234] Carl E Thoresen, PhD Carl E Thoresen and Michael J Mahoney. 1974. *Behavioral self-control*. Holt McDougal.

[235] Alice Thudt, Dominikus Baur, Samuel Huron and Sheelagh Carpendale. 2015. Visual mementos: Reflecting memories with personal data. *IEEE transactions on visualization and computer graphics*, 22 (1). 369-378.
[236] Alice Thudt, Bongshin Lee, Eun Kyoung Choe and Sheelagh Carpendale. 2017. Expanding research methods for a realistic understanding of personal visualization. *IEEE computer graphics and applications*, 37 (2). 12-18.
[237] Allison Tong, Peter Sainsbury and Jonathan Craig. 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International journal for quality in health care*, 19 (6). 349-357.
[238] Eric Topol. 2015. *The patient will see you now: the future of medicine is in your hands*. Basic Books.
[239] J Torous, JP Onnela and M Keshavan. 2017. New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. *Translational psychiatry*, 7 (3). e1053.
[240] Vincent W-S Tseng, Akane Sano, Dror Ben-Zeev, Rachel Brian, Andrew T Campbell, Marta Hauser, John M Kane, Emily A Scherer, Rui Wang and Weichen Wang. 2020. Using behavioral rhythms and multi-task learning to predict fine-grained symptoms of schizophrenia. *Scientific Reports*, 10 (1). 1-17.
[241] Amos Tversky and Daniel Kahneman. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5 (2). 207-232.
[242] Jenny Vafeiadou, Asimina Vasalou and George Roussos. 2021. Self-tracking in Parkinson's The Lived Efforts of Self-management. *Proceedings of the ACM on Human-Computer Interaction*, 5 (CSCW1). 1-25.
[243] Johan CH Van Bussel, Bernard Spitz and Koen Demyttenaere. 2006. Women’s Mental Health Before, During, and After Pregnancy: A Population-Based Controlled Cohort Study. *Birth*, 33 (4). 297-302.
[244] Bessel A van der Kolk, Laura Stone, Jennifer West, Alison Rhodes, David Emerson, Michael Suvak and Joseph Spinazzola. 2014. Original research yoga as an adjunctive treatment for posttraumatic stress disorder: A randomized controlled trial. *J Clin Psychiatry*, 75 (6). e559-e565.
[245] Signe Vikkelsø. 2005. Subtle redistribution of work, attention and risks: Electronic patient records and organisational consequences. *Scandinavian Journal of Information Systems*, 17 (1). 10.
[246] Stephen Voida, Mark Matthews, Saeed Abdullah, Mengxi Xi, Matthew Green, Won Jun Jang, Donald Hu, John Weinrich, Prashama Patil and Mashfiqui Rabbi. 2013. Moodrhythm: tracking and supporting daily rhythms. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, 67-70.
[247] Lauren S Wakschlag, Darius Tandon, Sheila Krogh-Jespersen, Amelie Petitclerc, Ashley Nielsen, Rhoozbeh Ghaffari, Leena Mithal, Michael Bass, Erin Ward and Jonathan Berken. 2020. Moving the dial on prenatal stress mechanisms of neurodevelopmental vulnerability to mental health problems: A personalized prevention proof of concept. *Developmental Psychobiology*.

[248] Roger Walsh. 2011. Lifestyle and mental health. *American Psychologist*, 66 (7). 579.

[249] Bradley V Watts, Brian Shiner, Lisa Zubkoff, Elizabeth Carpenter-Song, Julia M Ronconi and Craig M Coldwell. 2014. Implementation of evidence-based psychotherapies for posttraumatic stress disorder in VA specialty clinics. *Psychiatric Services*, 65 (5). 648-653.

[250] Peter West, Richard Giordano, Max Van Kleek and Nigel Shadbolt. 2016. The Quantified Patient in the Doctor's Office: Challenges & Opportunities. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, 3066-3078.

[251] Peter West, Max Van Kleek, Richard Giordano, Mark J Weal and Nigel Shadbolt. 2018. Common Barriers to the Use of Patient-Generated Data Across Clinical Settings. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, 484.

[252] Julia M Whealin, Emily C Jenchura, Ava C Wong and Donna M Zulman. 2016. How Veterans with post-traumatic stress disorder and comorbid health conditions utilize ehealth to manage their health care needs: a mixed-methods analysis. *Journal of medical Internet research*, 18 (10).

[253] Robin Whittemore, Susan K Chase and Carol Lynn Mandle. 2001. Validity in qualitative research. *Qualitative health research*, 11 (4). 522-537.

[254] Wendy Wood, Leona Tam and Melissa Guerrero Witt. 2005. Changing circumstances, disrupting habits. *Journal of personality and social psychology*, 88 (6). 918.

[255] Wanqing Wu, Sandeep Pirbhulal, Heye Zhang and Subhas Chandra Mukhopadhyay. 2018. Quantitative assessment for self-tracking of acute stress based on triangulation principle in a wearable sensor system. *IEEE journal of biomedical and health informatics*, 23 (2). 703-713.

[256] Tae-Jung Yun, Hee Young Jeong, Hee Rin Lee, Rosa I Arriaga and Gregory D Abowd. 2010. Assessing asthma management practices through in-home technology probes. In *2010 4th International Conference on Pervasive Computing Technologies for Healthcare*, IEEE, 1-9.

[257] Rachael Zehrung, Lily Huang, Bongshin Lee and Eun Kyoung Choe. 2021. Investigating Opportunities to Support Kids' Agency and Well-being: A Review of Kids' Wearables. *arXiv preprint arXiv:2104.05979*.

[258] Xuan Zhao, Niloufar Salehi, Sasha Naranjit, Sara Alwaalan, Stephen Voida and Dan Cosley. 2013. The many faces of Facebook: Experiencing social media as performance, exhibition, and personal archive. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 1-10.

[259] Haining Zhu, Joanna Colgan, Madhu Reddy and Eun Kyoung Choe. 2016. Sharing patient-generated data in clinical practices: an interview study. In *AMIA Annual Symposium Proceedings*, American Medical Informatics Association, 1303.