Toward Data-Driven Digital Therapeutics Analytics: Literature Review and Research Directions

Uichin Lee, Gyuwon Jung, Eun-Yeol Ma, Jin San Kim, Heepyung Kim, Jumabek Alikhanov, Youngtae Noh, and Heeyoung Kim

Abstract—With the advent of digital therapeutics (DTx), the development of software as a medical device (SaMD) for mobile and wearable devices has gained significant attention in recent years. Existing DTx evaluations, such as randomized clinical trials, mostly focus on verifying the effectiveness of DTx products. To acquire a deeper understanding of DTx engagement and behavioral adherence, beyond efficacy, a large amount of contextual and interaction data from mobile and wearable devices during field deployment would be required for analysis. In this work, the overall flow of the data-driven DTx analytics is reviewed to help researchers and practitioners to explore DTx datasets, to investigate contextual patterns associated with DTx usage, and to establish the (causal) relationship between DTx engagement and behavioral adherence. This review of the key components of data-driven analytics provides novel research directions in the analysis of mobile sensor and interaction datasets, which helps to iteratively improve the receptivity of existing DTx.

Index Terms—Causal inference, data-driven analytics framework, digital therapeutics (DTx), mobile and wearable data, technical and behavioral engagement.

I. INTRODUCTION

DIGITAL therapeutics (DTx), unlike traditional treatments such as pills, uses software installed in smartphones or wearable devices as software as a medical device (SaMD) to cure diseases and improve health conditions, which is a major departure from existing wellness products (e.g., Fitbits) [1]. As with traditional therapeutics, DTx also requires clinical validation of efficacy through systematic clinical trials [2].

The US FDA has already authorized a number of digital therapeutic products, for example, WellDoc's BlueStar [3] for diabetes management, and Pear Therapeutics' reSET [4] for

drug addiction recovery, opening up new DTx possibilities, such as doctors' prescriptions and insurance reimbursement. Unlike the traditional drug development, the cost of DTx development is relatively low, and new DTx markets are growing rapidly. The DTx Alliance, which was formed in 2017, consists of both startups (e.g., Omada Health [5] and Akili [6]) and global pharma (e.g., Novartis and Bayer). The DTx market is estimated to increase to $8.7 billion in 2025, with an average annual growth rate of 20% [7].

DTx therapies mostly consider behavioral changes in chronic diseases (e.g., diabetes and cardiovascular diseases) [8], [9] and neuropsychiatric diseases (e.g., depression, sleep disorders, and attention deficit hyperactivity disorder (ADHD)) [10]. These are the areas in which the treatment effects of cognitive behavior therapies are significant. DTx therapies can deliver patient-centered care by supplementing the areas in which treatment is difficult or poorly managed through existing treatment methods (e.g., lifestyle coaching and cognitive behavior therapies), to improve the quality of care at lower costs.

One of the important components of traditional drug development is the selection and optimization of a drug delivery system that aims to effectively deliver a specific drug to the desired target (e.g., sustained release with microneedle patches) [11]. Digital therapeutics can deliver various interventions through digital technologies (e.g., interactive mobile content, videos, chatbots, and push notifications) [12], [13]. Thus, it is very important to analyze and optimize the engagement and receptivity of "DTx delivery systems" using mobile and wearable devices.

The existing drug delivery systems can be evaluated in controlled environments. However, DTx usage is a daily occurrence in the lives of patients, and thus, it is very difficult to evaluate the real-world user experiences and the efficacy of DTx in a laboratory setting [13]. Traditional clinical trials on DTx mostly focus on measuring the endpoints or proximal/ distal outcomes in the wild, but less attention has been paid to systematically understanding DTx user engagement and adherence patterns, which are essential for DTx improvement. Furthermore, there is a lack of agreement on the methods and criteria for evaluating DTx related user experiences [14].

The receptivity to DTx relates to the overall process of intervention delivery using digital devices (e.g., notification delivery, notification perception/checking, and behavioral adherence) [15]. DTx aims to induce behavioral changes in users, and it is very important to analyze patient DTx engage-
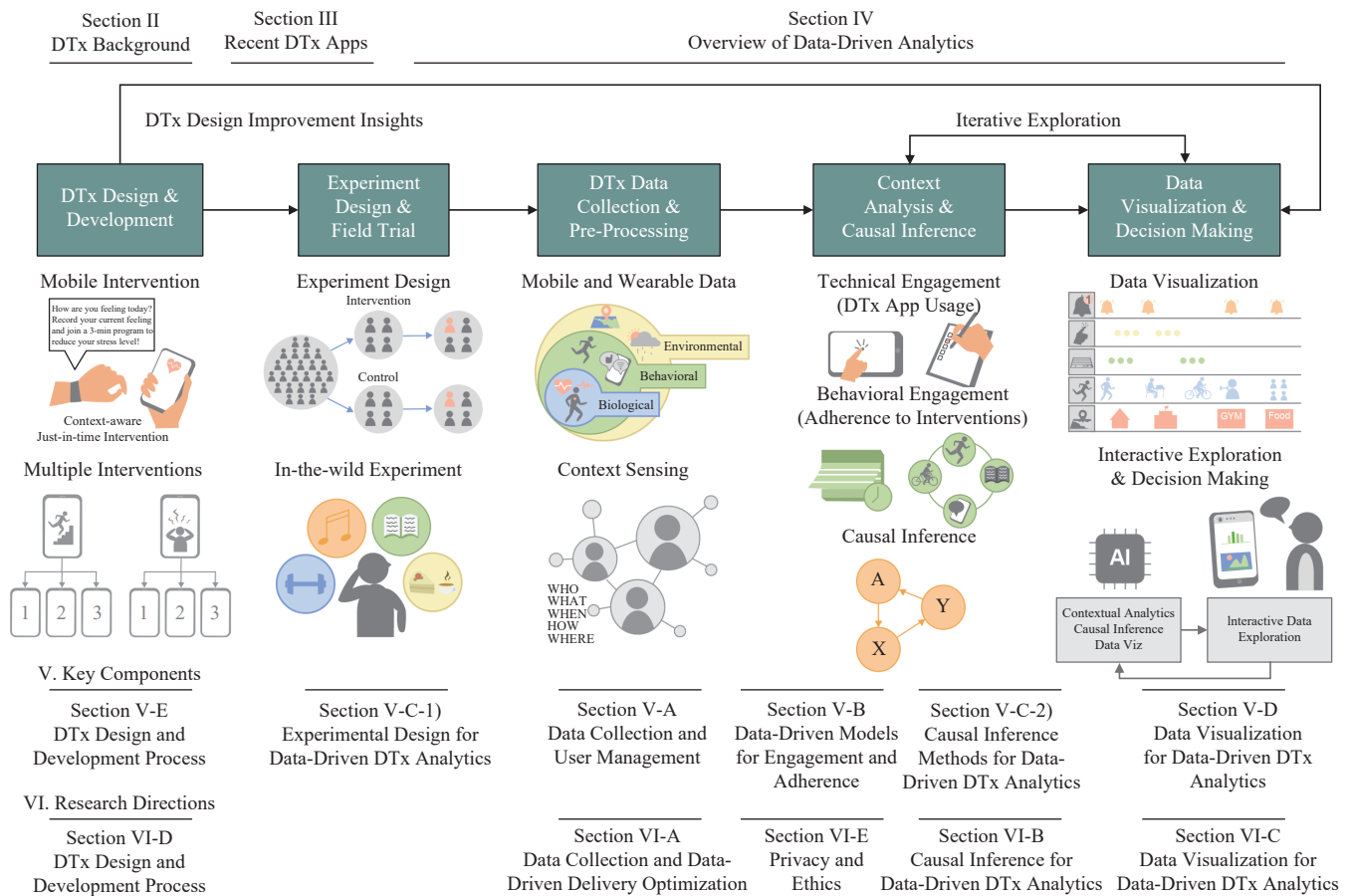
Fig. 1. A conceptual diagram of data-driven DTx analytics: a process overview (Section IV) and key component reviews (Section V).

ment and receptivity to shorten the DTx development time and to maximize the effectiveness of DTx. This review aims to illustrate the flow and key components of data-driven DTx analytics to help researchers and practitioners to investigate the user engagement and intervention receptivity to DTx by analyzing digital footprint data (known as digital phenotype data) collected from mobile and wearable devices. Data-driven DTx analytics for user engagement and intervention receptivity in DTx delivery systems will provide key insights for the improvements of DTx, possibly innovating the existing paradigm of DTx development processes.

Data-driven DTx analytics is closely related to automation research because DTx replaces human-based health interventions with mobile-based counterparts. Furthermore, intelligent agents leverage context sensing and data processing to automate and personalize health services. This topic is also related to human-in-the-loop system design for proactive guiding of user behavior where an operator's behavior and machine intelligence act as human-cyber-physical systems that are critical for automation performance [16]. Automation fields have a long tradition of using sensor data to optimize the performance of human-machine collaboration. This work on data-driven DTx analytics broadens the scope of existing automation research.

The remainder of this paper is organized as follows (Fig. 1). In Section II, background information (e.g., DTx definition and regulations) is provided. In Section III, an overview of recent DTx is provided, and opportunities for adopting just-in-time interventions in DTx are discussed. In Sections IV and V, an overview of the flow of data-driven DTx analytics and a detailed review of key components are provided. In Section VI, to conclude, several directions for future research are discussed.

II. DTx BACKGROUND

A. Defining DTx and its Relationship to Digital Health

The Digital Therapeutics Alliance, an industry association in the digital health area, defines DTx as "evidence-based therapeutic interventions to patients that are driven by high-quality software programs to prevent, manage, or treat a medical disorder or disease" [2]. In academia, DTx is similarly defined, e.g., "a new treatment modality in which digital systems (e.g., smartphone apps) are used as regulatory body-approved, prescribed therapeutic interventions to treat medical conditions" [1]. Existing DTx mostly target chronic diseases with continuous intervention support to change behaviors or lifestyles and to help patients manage their health conditions effectively. The coverage of therapeutics includes obesity, prediabetes, hypertension, hyperlipidemia, smoking-related diseases, chronic pain, chronic obstructive pulmonary disease and asthma, diabetes, alcoholism, coronary artery disease, and serious mental illnesses [17], [18].

Although DTx has recently been in the spotlight, the use of digital products in providing therapeutic interventions has

been studied for a long time. As Webb *et al.* [19] state in their study, researchers have designed therapeutic interventions on theoretical bases, applied behavioral change techniques in them, and delivered them via the Internet. The term "digital therapeutics" was first mentioned in a study published in 2015 [20], defining it as "evidence-based behavioral treatments delivered online that can increase accessibility and effectiveness of health care." As this definition implies, it became important to verify the effectiveness of DTx treatments (i.e., the actual outcome of the DTx treatment in the real world) [21], [22]. Additionally, the industry has also begun to set up standards for clinical evidence.

Digital therapeutics is deeply related to digital health and mHealth and is generally classified as a subset of digital health. According to the World Health Organization, the origin of digital health is eHealth, which is defined as "the use of information and communications technology in support of health and health-related fields." This definition includes mHealth, which is "the use of mobile wireless technologies for health." As a subset of eHealth, mHealth expands its scope to emerging technologies, such as big data analysis, artificial intelligence, and genomics [23]. The US FDA, on the other hand, extends the area of digital health to cover "mobile health (mHealth), health information technology, wearable devices, telehealth, telemedicine, and personalized medicine" [24]. There is another subset of digital health named "digital medicine," which is defined as "high-quality hardware and software that support the practice of medicine broadly" for measurement and intervention in health-related services [25], [26]. For instance, digital medicine may include pills with built-in sensors for monitoring tumors inside the body or wearables that continuously track the glucose level or heart rate. Healthcare providers or consumers then utilize the collected data to manage the health status by changing treatment decisions or medication doses. In this sense, digital medicine can be seen as a broader concept than DTx [27]. Compared with digital medicine, most DTx products focus only on software and emphasize making direct changes in health conditions or adjusting treatments based on the collected health data.

Digital health uses technology to deliver information and enable communication, with the purpose of monitoring and managing patients and consequently improving their health conditions [28]. Moreover, this approach reduces the burden of health care, allows patients to manage their health even outside traditional hospitals, and individualizes the treatment by implementing behavior change theory or techniques [9]. Following these definitions and descriptions, digital therapeutics can be considered part of digital health. In particular, during the COVID-19 era, the US government relaxed the regulations for noninvasive remote monitoring devices and reduced direct contact between patients and healthcare providers [29]. This is expected to increase the use of DTx in everyday life and eventually address health inequalities by supporting patients who cannot easily access medical facilities, such as people living in rural areas [30].

Various studies on DTx have been conducted in the area of digital health using different digital platforms (e.g., web and mobile apps). Digital health research targets many diseases or health issues, such as diabetes [8], cardiovascular diseases [9], asthma [31], mental health (e.g., depression, anxiety, ADHD, autism spectrum disorders, and eating disorder) [10], weight management [32], and smoking cessation [33]. In addition, digital health solutions could be utilized to manage various types of cancer tumors [34], support patients in rehabilitation from neurological diseases (e.g., stroke, Parkinson's disease, and multiple sclerosis) [35], [36], and prevent and treat HIV [37], [38]. A report published in 2017 showed that there are more than 318,000 digital health apps publicly available to consumers through Apple Store and Google Play, with more than 200 apps added each day [39].

However, not all apps can be classified as "digital therapeutics" because clinical evidence and real-world outcomes (e.g., managing, preventing, or treating a medical disorder, or optimizing medication) have to be satisfied for regulatory purposes. Therapeutic interventions must also be certified by regulatory bodies in terms of efficacy and safety for intended use [26]; by demonstrating clinical evidence (or efficacy) via randomized controlled trials (RCTs) in which participants are randomly assigned to the clinical interventions or the control group set up for comparison; the control could be a placebo or no intervention at all [39]. Research communities on digital health have been striving to conduct RCTs to obtain clinical evidence. However, the generation of clinical evidence for efficacy and safety is still scarce because of the time and cost of running large-scale RCTs, which has become one of the main challenges in DTx. Further issues might arise such as the technology becoming outdated during the trial, ethical considerations, infeasible placebo control groups, or privacy concerns arising from remote informed consent [1].

To summarize, the relationship between digital health, mHealth, and DTx can be described as shown in Fig. 2. Both mHealth and DTx are subsets of digital health and overlap in some areas.
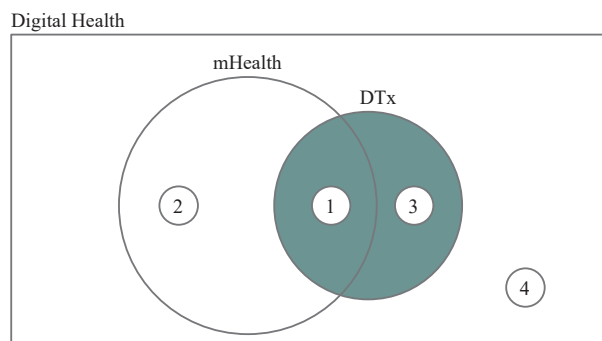
Fig. 2. Relation among Digital Health, mHealth, and DTx. There are four different categories in Digital Health depending on whether the system utilizes mobile devices (mHealth) and whether it has sufficient direct therapeutic intervention or clinical evidence (DTx).

1) mHealth and DTx: DTx products that deliver therapeutic interventions via mobile devices are included in this area. Most mobile applications and major DTx products use this format.

2) mHealth but not DTx: Digital health apps are available on

mobile devices (usually smartphones), but they do not demonstrate clinical evidence or real-world outcomes.

3) DTx but not mHealth: DTx products that do not utilize mobile or wearable platforms. These products may be in the form of software applications or web applications.

4) Digital health, neither mHealth nor DTx: Digital health apps available as a form of software application or web application (not smartphones), but have only the ability to capture, store, display, or transmit health data and information with no direct therapeutic interventions, clinical evidence, or real-world outcomes (e.g., health information technology, telehealth, and medical imaging)

B. DTx Regulations

The development of the digital healthcare sector has led to the emergence of new software concepts such as software as a medical device (SaMD). The International Medical Device Regulator Forum, which is a coalition of regulators from various countries, defines SaMD as "software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device" [57]. In general, SaMD includes DTx as well as other clinical decision-making software such as computer-aided diagnosis for radiologists. Similarly to hardware-based medical devices, SaMDs have highlighted the need for regulatory oversight and approvals. the US FDA requires SaMD to receive approval by demonstrating safety and effectiveness either through the Premarket Notification 510 (K) process of comparison with *similar* legally marketed devices (known as substantial equivalence proof) [58], or the De Novo classification process if a comparison is not feasible (new technology) [59]. In 2018, US FDA further introduced breakthrough designations for SaMDs with more effective treatment or diagnosis of life-threatening human diseases such that manufacturers can interact with the FDA's experts to expedite such review processes [60].

There is also a newly created pre-certification program that certifies SaMD "developers" instead of SaMD "products" so that the total product lifecycle can be managed (e.g., product quality, patient safety, and cybersecurity) [61].

To date, most applications and services, classified as DTx, have received 510 (K) clearance by submitting RCT results from users with new technology and those with make-believe technology as a placebo to demonstrate their effectiveness. WellDoc's BlueStar diabetes management system was one of the first FDA 510 (K) submissions approved in January 2017. Pear Therapeutics reSET (drug addiction) and Akili's EndeavorRx (attention deficit) were approved via the De Novo process and validated by comparing existing treatment methods with and without the newly developed techniques [6], [62].

Besides the US FDA approvals, for diabetes management, there is another certification called the full Centers for disease control and prevention (CDC) recognition for diabetes prevention program (DPP). Several commercial DTx products have received full CDC recognition for DPP, such as Noom [63] and Omada [64], which provide users with diabetes programs based on computerized cognitive behavioral therapy. They were recognized for the effectiveness of diabetes management through clinical trials for participants in the National

DPP program, which lasted more than 16 weeks [5], [65].

III. REVIEW OF RECENT DTX THERAPIES

Digital therapeutics apps use a variety of behavior change techniques such as feedback on behavior and goal setting, social support, instructional guidelines, and self-monitoring of behavior, and outcome(s) of the behavior. As DTx apps are not yet well established, to grasp the trends, a non-probability sampling method---snowball sampling---was performed. Potential DTx apps were identified online based on our judgment of mental and physical disorder categories until the required sample size was reached. A solid effort was made to embrace most DTx apps introduced/released in the last three years, as well as those with FDA certificates (i.e., De Novo and 510 (K)) and recognition by the CDC. This paper may not exhaustively review all relevant literature, but it clearly reflects the state-of-the-art technology of DTx apps. The insights obtained from the process are summarized according to the therapeutic areas and diseases they target. For each DTx, regulatory endorsement is checked (e.g., 510 (K), De Novo, and CDC full recognition), and intended use is categorized based on Digital Therapeutics Alliance [2] categories: 1) management or prevention of a medical disorder/disease, 2) medication optimization, and 3) medical disease or disorder treatment.

As shown in Table I, the established DTx apps span two therapeutic areas, namely, mental and physical disorders. The mental disorders include ADHD [6], depression [40], [41], insomnia [42], panic disorder [43], and substance abuse [4], [45]. Among DTx apps in this area, Pear Therapeutics's reSET [4] is the first prescription DTx product (i.e., De Novo) designed to deliver behavioral therapy to treat substance use disorder; reSET-O [45] is another prescription DTx product designed to deliver cognitive behavioral therapy for opioid use disorder. As a research prototype, EndeavorRx [6] is a home-based, video game-like digital app for the treatment of inattention and cognitive dysfunction in pediatric patients with ADHD.

Another major branch of DTx apps is physical disorders, such as asthma [46], cancer [48], diabetes [3], [49], [51], diabetes/hypertensions [52], [5], [53], and diabetes/obesity [50]. Among the DTx apps in this area, BlueStar [3], a digital therapeutic app developed by WellDoc, has assisted patients and providers in improving glucose control by using real-time data and feedback to support healthy behaviors, such as medication adherence, diet and exercise control, and psychosocial wellness. BlueStar has demonstrated the capacity to shift HbA1c levels in populations with diabetes. Interestingly, the Dario Blood Glucose Monitoring System [69] by DarioHealth, Omada [70] by Omada Health and Livongo [71] by Teladoc Health combine the DPP with hypertension, as 80% of people with type 2 diabetes also have high blood pressure. Both hypertension and diabetes result from metabolic syndromes. Thus, they develop sequentially in the same individual [72], [73].

DTx apps can incorporate *just-in-time support* which is an attempt to provide the right type (or amount) of support at the right time (i.e., neither too early nor too late). For example,

TABLE I
RECENT DIGITAL THERAPEUTICS

Category	Disorder/Disease	DTx name	Company	Status	Intention
Mental disorder	ADHD	EndeavorRx [6]	Akili Interactive Labs	FDA De Novo	Treat
	Depression	deprexis [40]	GAIA-AG	FDA 510 (K)	Manage
		Woebot [41]	Woebot Health	−	Manage
	Insomnia	Somryst [42]	Pear Therapeutics	FDA 510 (K)	Manage
	Panic disorder	FREESPIRA [43]	Freespira	FDA 510 (K)	Manage
		NightWare [44]	NightWare	FDA De Novo	Manage
	Substance abuse	reSET [4]	Pear Therapeutics	FDA De Novo	Treat
		reSET-O [45]	Pear Therapeutics	FDA De Novo	Treat
Physical disorder	Asthma	Propeller [46]	Propeller Health	FDA 510 (K)	Optimize
		NuvoAir [47]	NuvoAir	FDA 510 (K)	Manage
	Cancer	Oleena [48]	Voluntis	FDA 510 (K)	Manage
	Diabetes	BlueStar [3]	WellDoc	FDA 510 (K)	Manage
		Insulia [49]	Voluntis	FDA 510 (K)	Manage
		Transform [50]	Virgin Pulse	CDC DPP	Manage
		Virta [51]	Virta Health	−	Manage
		Dario [52]	DarioHealth	FDA 510 (K)	Manage
	Diabetes, Hypertension	Omada [5]	Omada Health	CDC DPP	Manage
		Livongo [53]	Teladoc Health	FDA 510 (K)	Manage
	HIV	PositiveLinks [38]	PositiveLinks Platform	−	Manage
	Stroke, Neurological disorders	Constant Therapy [54]	Constant Therapy Health	FDA Breakthrough	Manage
		NeuroEyeCoach [55]	NovaVision	FDA 510 (K)	Manage
		Nirvana [56]	BTS Bioengineering	−	Manage

the ongoing monitoring of an individual with mobile sensing can identify when these events/conditions occur (i.e., when support is needed). As an emerging technology-driven, behavior change-oriented intervention type, just-in-time adaptive intervention (JITAI) capitalizes on real-time sensor data collected via mobile sensing technology (e.g., smartphones and wearables) to adaptively trigger appropriate support in situ [74]. The major components of DTx with JITAI can be summarized as follows:

1) Distal Outcome: The long-term goal (i.e., behavior change) of a DTx app (e.g., reduction of sedentary behavior in older adults [75]).

2) Proximal Outcome: The short-term goal of a DTx product, which mediates the effect of the intervention on the distal outcome. In other words, the proximal outcome ensures progress toward the distal outcome (e.g., increased number of active breaks from prolonged sitting over a day).

3) Tailoring Variables: The collection of behavioral and contextual data that can identify when behavioral support might be most effective (e.g., accumulated sitting time, location of individual, time of day, response to support prompts sent earlier, and frequency of support prompts).

4) Intervention Options: Adequate intervention options (e.g., suggestions of light movements, positive feedback, and encouragements to repeat the light movements) that are delivered to the user when an opportune moment for behavior change support is detected.

5) Decision Points: A marked moment when a decision to send an intervention option is made considering tailoring variables (e.g., the time between 5 pm and 9 pm).

6) Decision Rules: Systematic rules of whether to provide certain intervention options based on past intervention options and tailoring variables (e.g., location-based feedback is provided if a user becomes sedentary for more than 30 minutes).

JITAI-based DTx apps were further reviewed in the context of JITAI with relevant data types in Table II, which describes DTx apps that target alohol addiction [66], tobacco addiction [67], and obesity [68].

Alcohol Addiction: A-CHESS [66] offers JITAIs to reduce risky drinking days (i.e., distal outcome) for patients with alcohol use disorder after leaving a treatment facility. A-CHESS monitors data collected via mobile sensing to trigger just-in-time interventions in the form of messages and tracks the reduced brief alcohol monitoring index (i.e., proximal outcomes) weekly to make progress toward the distal outcome. The app uses tailoring variables (e.g., location of an individual, user data submissions, messages sent/received via the app) and intervention options such as sending a prompt notice if the user comes near places associated with a high risk for drinking. A-CHESS defines the moment when the user's location is close to the high-drinking-risk place and the moment when the panic button is pressed as decision rules and decision points.

Cigarette Smoking: JITAI for smoking cessation can leverage self-reports and automatic sensing for risk assessment. Smart-T [76], [77] estimated the risk of smoking using the

TABLE II
DIGITAL THERAPEUTICS WITH JITAI

App name (target disease)	JITAI components	Contents	Data types
A-CHESS (alcoholism) [66]	Distal outcome	Regular reduction on the number of risky drinking days (4, 8, and 12 months after discharge)	numeric
	Proximal outcome	Weekly Brief Alcohol Monitoring (BAM) index result	numeric
	Tailoring variables	Location of the individual	object
	Intervention options	Helping patients stay sober by prompting them if they go near the high risk places	string
	Decision points	The moment when the user gets closer to high-risk location	timestamp
		When the user presses the panic button	timestamp
	Decision rules	Whether the user gets closer to the high-risk location or not	boolean
		Whether the panic button is pressed or not	boolean
Q-Sense (smoking) [67]	Distal outcome	Smoking abstinence	boolean
	Proximal outcome	Reduced stays within high risk locations	numeric
	Tailoring variables	Location of the individual	object
		User self-reports (i.e., mood, stress, urge, current context, and whether other smokers are present) collected when the user starts smoking during the pre-quit phase	mixed
	Intervention options	Tailored support message (prefilled based on the user's demographics and smoking survey)	string
		Feedback messages based on the user's smoking reports	string
	Decision points	When the user enters and stays in the high risk location for more than 5 minutes	timestamp
	Decision rules	Whether the user enters and stays in the high risk location for more than 5 minutes	boolean
DietAlert (obesity) [68]	Distal outcome	Weight loss	numeric
	Proximal outcome	If exceeding calorie limit or not	boolean
	Tailoring variables	21 tailoring variables (temptation, boredom, hunger, and planning food intake, exercise, etc.)	numeric
	Intervention options	157 subdivided elements provided (e.g., think of something fun to try right now)	string
	Decision points	When the user replies to an EMA prompt	timestamp
	Decision rules	Use supervised machine learning to identify risk for lapse occurrence	numeric

user's self-report data via ecological momentary assessment (EMA) (e.g., urge to smoke, stress, alcohol consumption, and availability of cigarettes). Sense2Stop [78] leverages automatic stress sensing mechanisms to deliver support messages when high-stress moments are detected. Q-Sense [67] learns a user's smoking behavior, emotional conditions, and places via EMA, and the user's high-risk locations are tracked to deliver just-in-time support messages. Q-Sense leverages user self-assessment and mobile sensing to trigger just-in-time support messages for smoking abstinence (i.e., distal outcome). It also tracks proximal outcomes (i.e., not staying in a high-risk location) to make progress toward the distal outcome. Self-assessment and location are used as tailoring variables to trigger the intervention. Q-Sense checks whether a user stays in high-risk locations for longer than 5 min (known as a geofensing technique), which is considered a decision rule, and the current timestamp becomes the decision point. Intervention options include tailored support and feedback messages. Tailored support messages are pre-populated based on the user's demographics and smoking survey. Feedback messages are sent based on the smoking behavior pattern of the user and on user self-reports (e.g.,"Based on 12 reports, did you know you smoke 25% of the time when working?").

Obesity: DietAlert [68] focuses on lapses in a weight-control diet among overweight and obese individuals. For just-in-time intervention, the app tracks whether the user exceeds the calorie limit as the proximal outcome. DietAlert tracks 21 tailoring variables (e.g., temptation, boredom, hunger, and exercise) and 157 subdivided elements (e.g., thinking of something fun for boredom). DietAlert uses supervised machine learning to identify lapse occurrence risk as a decision rule and keeps track of the user responses for decision points (six semi-random time intervals throughout the day, spaced approximately 2–3 h apart).

IV. OVERVIEW OF DATA-DRIVEN DTX ANALYTICS

This section provides an overview of the core elements of data-driven DTx analytics (Fig. 3). The major flow of data-driven DTx analytics involves DTx user experience, DTx usage, and behavior adherence (proximal and distal outcomes). These elements are closely related to the offered DTx mechanisms (e.g., goals and strategies) and users' intervention contexts (e.g., user traits and environments). In addition to traditional clinical trials with RCTs, we discuss how sensor and interaction data can be used to enable contextual and causal analytics.

DTx Engagement: We used the existing conceptual model *engagement* by Yardley *et al.* [79], which made distinctions between engagement with technological and behavioral aspects. Engagement with technological aspects (i.e., DTx user experience and DTx usage) refers to how a user makes use of software for behavior changes. In contrast, engagement with behavioral aspects refers to how a user initiates and sustains behavior changes, which can possibly lead to "positive
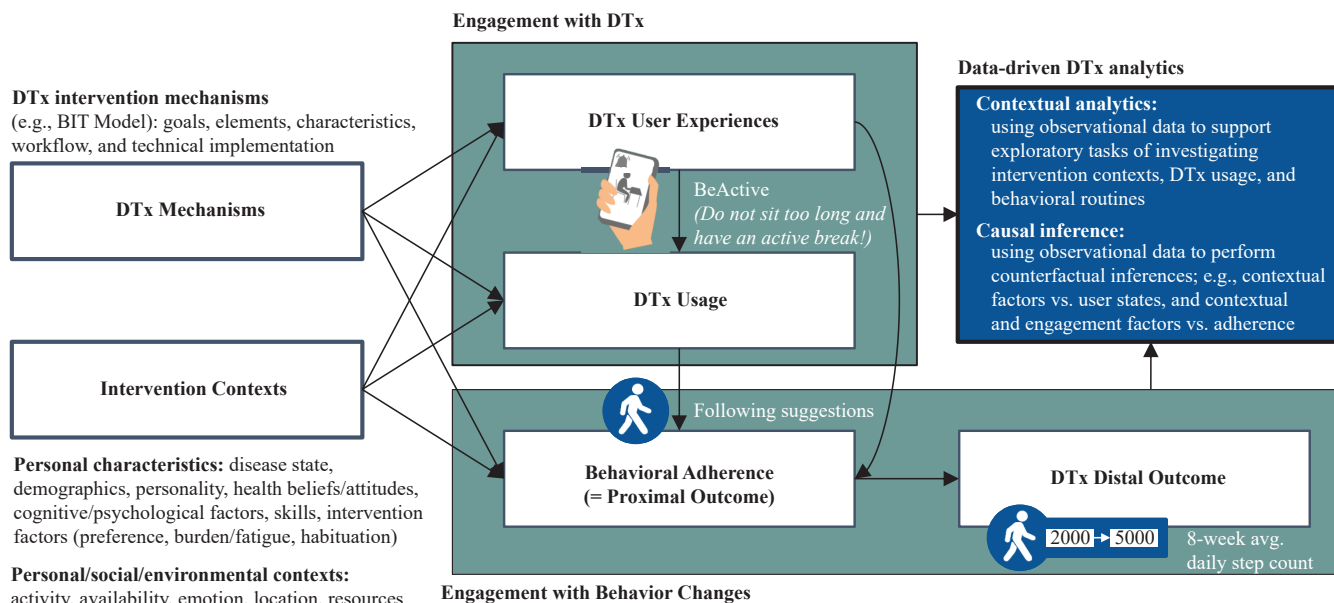
Fig. 3.   Data-driven DTx analytics framework.

outcomes" (e.g., reaching weight loss goals). Intervention software plays an important role in scaffolding behavior changes, so that sustained usage of intervention software is no longer necessary for achieving "positive outcomes." In other words, if a user has finished mastering a skill offered by the intervention software, behavioral changes can be successfully maintained even without software usage. This distinction was also used in Alshurafa *et al.*'s work [80], where engagement with intervention software referred to DTx usage (e.g., duration and frequency of app usage) and engagement with behavior changes referred to *adherence to the intervention* (e.g., proximal outcomes). In this work, "engagement" refers to DTx software usage and user experience, and "adherence" refers to behavior changes prompted by intervention. In Section V-B, the modeling of engagement and adherence using datasets is further discussed.

DTx Mechanisms: Existing intervention models include the behavior intervention technology (BIT) model [12] and Fogg's model [81]. The BIT model is a software-based intervention at both theoretical and technical levels: 1) theoretical aspects include intervention goals (e.g., weight loss and healthy eating) and behavioral intervention strategies (e.g., goal setting, education, monitoring, and feedback), and 2) technical aspects include behavioral intervention technology elements (e.g., information delivery, data collection, and reports), technical characteristics (e.g., medium, complexity, aesthetics, and personalization), and workflow of the intervention to determine when to deliver the intervention to the user (e.g., context-based scheduling).

Intervention Contexts: These include both personal characteristics and contextual factors. An existing behavior change model for Internet interventions [82] considers 1) user characteristics (e.g., disease states, demographics, traits, beliefs and attitudes, physiological factors, and skills), 2) intervention experiences (e.g., user preferences, perceived burdens/fatigue, and habituation), and 3) intervention environments (e.g., personal, professional, and community aspects). This "contex-

tual" model can be further extended by using existing context representation models for ubiquitous computing where mobile, wearable, and Internet-of-Things (IoT) devices are used to enable context-aware services [83], [84]. Additional contextual information includes both human factors (e.g., users' cognitive and affective states, social environment, and tasks) and physical environments (e.g., location/place, lighting condition, and temperature). Intervention context data can be collected via user self-reports and automatic inferences from sensor data collected passively via machine learning.

Data-Driven DTx Analytics: An RCT, as the gold standard for clinical evaluation, helps to evaluate the efficacy of a newly developed digital intervention, as opposed to the traditional approach. However, it is difficult for RCTs to identify the key contextual patterns and intervention factors that influence efficacy. Sensor and interaction data passively collected during RCTs can be used in contextual and causal analytics. Contextual analytics supports exploratory tasks of understanding contextual factors related to DTx usage such as lifestyle (behavioral routines), intervention contexts, and psychological states of the user (e.g., emotion and stress) over the intervention period. Causal analytics (counterfactual inference) helps to investigate how individual intervention elements and user engagement influence behavioral adherence (i.e., proximal and distal outcomes).

V. KEY COMPONENTS OF DATA-DRIVEN DTX ANALYTICS FLOW

The overall flow of data-driven DTx analytics is shown in Fig. 1. Conducting large-scale RCTs is very expensive and time-consuming and constitutes a major bottleneck in rapid software iteration. The goal of data-driven DTx analytics is not only to evaluate the efficacy of the intervention, but also to acquire design and clinical insights through a comprehensive analysis of the passively collected sensor and interaction data from DTx field trials. The major components of a DTx analysis flow include (1) data collection and user manage-

TABLE III
COMPARING MAJOR FEATURES OF EXISTING CTMS SYSTEMS

CTMS	Major features supported							
	Milestone mgmt	Education mgmt	Resource mgmt	Subject mgmt	Data collection	Data analytics	Legal service	Financial service
ArisGlobal [85]	O	O	O	O	O	O	O	O
BioClinica [86]	O	O	O	O	O	O	O	O
Bio-Optronics [87]	O	O	O	O	O		O	O
DSG [88]	O	O	O	O	O		O	O
eResearch Technology [86]	O	O	O	O	O	O	O	
MedNet Solutions [89]	O	O	O	O	O	O	O	O
Medidata Solutions [90]	O	O	O	O	O	O	O	O
PAREXEL International [91]	O	O	O	O	O	O	O	O
Trial By Fire Solutions [92]	O			O	O		O	O
Veeva Systems [93]	O	O		O	O	O	O	O

O: Supported

ment in field trials, (2) data-driven models for engagement and adherence, (3) experimental design and contextual/causal analytics, (4) data visualization, and (5) DTx design and development processes, which are reviewed in detail as follows.

A. Data Collection and User Management

We review technologies enabling data collection and user management, such as clinical trial management systems and general-purpose platforms. Data sources are largely based on sensor and interaction data from smart devices (e.g., phones, wearables, and IoT devices); thus, smart devices and their interactions are discussed.

1) Clinical Trial Management System: As stakeholders perform various clinical trials with a certain level of quality, it is necessary to utilize an information technology system to support data collection and user management in standardized clinical trial processes, to comply with relevant regulations. Based on a commonly agreed upon standardized process, a clinical trial management system (CTMS) was designed and developed as a comprehensive system to comply with privacy and security regulations. Table III summarizes the representative CTMSs [94] and states the eight key supporting features required to conduct standardized clinical trials, as discussed in [95]. These features include a) milestone management (e.g., managing timeline of clinical trial such as regulatory completion), b) education management (e.g., training of researchers of the study by CTMS expertise team or providing educational services to ensure that subject remains in the study), c) resource management (e.g., managing medical devices and clinical drug import and export), d) subject management (e.g., managing the health information of the subject and subject recruitment and scheduling), e) data collection (e.g., using electronic data capture (EDC) system to collect clinical data in electronic format for human clinical trials instead of the traditional paper-based data collection methodology, to streamline data collection, and to expedite the time for marketing the drugs and medical devices), f) data analytics (e.g., providing visualization of clinical research data using dashboard or ana-

lyzing clinical data using other analytic modules), g) regulatory services (e.g., adhering to guidance and directives from the regulatory agencies such as FDA and European Medicines Agency), and h) financial services (e.g., financial management of clinical trials).

Ten well-known CTMSs reported in a recent study [94] were selected for this review. Product brochures, publicly available videos, and white papers were scrutinized to check whether the eight essential features were supported. Most systems fully support the key features; however, several systems do not provide some of the features, possibly because their focus is on specific customer segments. It is interesting to note that Medidata Solutions [90] and Parexel International [91] provide additional EDC support to collect sensor data (e.g., pulse oximeters, blood pressure meters, activity trackers, glucometers, spirometers, body weight scales, and ECG) and wearables. The other CTMSs do not closely match these standards. Major players include IBM Watson Health [96], Oracle [97], Nextrials [98], and Winchester Business Systems [99].

2) General-Purpose Mobile Sensing Platforms: Compared with CTMS, research communities in mobile and ubiquitous computing have made intensive efforts to build general-purpose platforms to collect sensor data from mobile and wearable devices (e.g., smartwatches and smartbands). As shown in Fig. 4 , here, the platform components are conceptualized with a layered architecture. a) *Data source and Communication interfaces* layer includes the collection of sensor data (e.g., accelerometers, gyroscopes, magnetometers, and GPS) from smartphones and other wirelessly connected devices (e.g., smartwatches, smartbands, and IoT devices). Through a communication interface (i.e., Bluetooth, WiFi), data from the devices are transferred to smartphones and stored locally. b) *Data acquisition* layer includes the implementation of an agent that performs basic functions (e.g., device registration, connection, subscription, and data queries), sampling rate adjustment for sensing, and connection failure management (e.g., wireless connectivity failures between a wearable and smartphone). The obtained data are then archived into a local
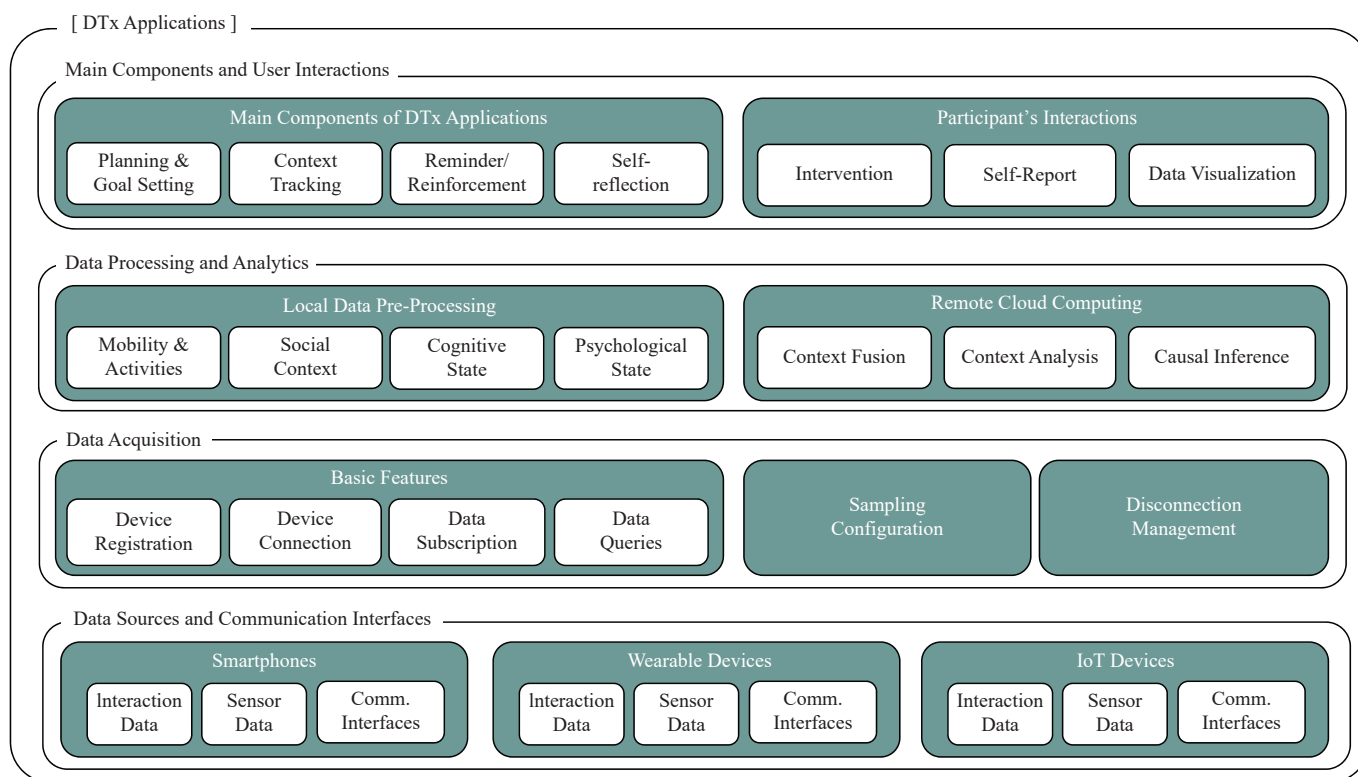
Fig. 4.   The architecture of data-driven DTx analytics platform.

database. c) *Data processing and analytics* layer implements local data preprocessing and remote cloud computing for context analytics and causal inference. Various machine learning algorithms can be executed on either local devices (e.g., smartphones or local servers) or remote servers. Lightweight machine learning algorithms on edge devices process raw sensor data to extract contextual features and behavioral markers (e.g., mobility features, physical activities, and users' cognitive and psychological states). Computationally intensive analytics can be performed via remote cloud computing infrastructure. For example, causal inference helps to establish causality of the effect of the provided DTx treatments, and contextual analytics helps to understand users' intervention contexts, such as problematic situations and the influence of user contexts on adherence. d) *DTx* layer includes the main components of DTx applications and user interactions. DTx applications (e.g., smoking cessation and depression management) usually contain pipe-lined procedures (i.e., planning and goal setting, context tracking, reminding/reinforcement, and self-reflection). The participant's interactions involve behavioral interventions, user self-reporting to the platform, and visualization of the user's daily (or longer-term) summary toward the goal. A more detailed review on DTx mechanisms can be found in Section V-E.

General-purpose platforms mainly focus on providing reliable and scalable sensor data collection, and therefore, compared to CTMS, offer only a limited number of functionalities, such as research resource/progress management and logistics support. Recently, several well-known data-collection platforms have been introduced, including OpenDataKit [100], AWARE [101], and mCerebrum [102], which enable resear-

chers to collect user self-reports and stream high-frequency sensor data to the remote cloud for storage and data processing. OpenDataKit adopts a middleware design approach that allows third-party app developers to minimize their efforts in sensor-specific codes via reusable sensor drivers. For example, it is possible to download new sensor capabilities from the application market and use them without any need for modifications. OpenDataKit further manages the discovery, communication channels, and data buffers for extensibility. AWARE provides mobile data-logging tools and supports external sensor plugins for the collection and abstraction of sensor data for context-aware service delivery. System scalability is an important issue for these platforms. mCerebrum significantly improves the scalability of storage for high-rate sensor data and provides several fine-tuned features, such as sensor duty cycling, energy-optimized context inference computation as a shared service, and sensor data quality assessment. AWARE and mCerebrum can deliver local interventions, but the cloud components of these frameworks are limited to automatically triggered interventions based on past user data. Data collection requires sustained user engagement. SARA [103] integrates gamified engagement strategies, including contingent rewards, badges for completing active health tasks, funny memes/gifs and life insights, and health-related reminders or notifications to incentivize data collection.

There are other data collection platforms such as the Personal Health Intervention Toolkit (PHIT) [104] that exclusively target health interventions. PHIT is a framework that facilitates mobile data collection and supports health interventions for research purposes. PHIT allows researchers to

explore and perform health status analysis, intervention recommendations, and self-help activities, and to explore uploaded data through a Web-based portal. The major departure of PHIT from CTMS and general-purpose data collection platforms is that it provides an intelligent virtual advisor that analyzes real-time data from devices and facilitates tailored interventions. A recent study proposed Duty [105], [106], which integrates mindfulness-based relaxation, behavioral education in sleep quality and alcohol use, and psychometric and psychophysiological data capture. Duty used PHIT as a personalized health intervention framework for acquiring data, including self-report instruments, EMA diaries, cognitive tests, and game-like activities.

3) Smart Devices and User Interactions: Current commercial DTx apps (e.g., Woebot and reSET) are mostly in software format using off-the-shelf smartphones, but they also leverage wearable trackers (e.g., Fitbit and Apple Watch) and IoT devices (e.g., Propeller Health's Bluetooth inhalers and Insulia's Blood Glucose Meters). Mobile, wearable, and IoT technologies support fine-grained sensing and tracking of users' states ranging from physiological signals, such as heart rates and skin temperature, to physical activities, social interactions, and user's interactions with digital devices (also known as human-computer or human-machine interactions). These smart devices provide *physical actuation* (e.g., controlling light bulbs or thermometers) or *virtual actuation* (e.g., launching apps and sending emails). Beyond local physical sensing, it is possible to use Web-based sensor data such as weather and air quality data via open application programming interfaces (APIs) as virtual sensors [107].

Existing DTx apps that consider just-in-time support often use diverse real-time sensing features such as activity tracking. For example, the BeActive system continuously monitors a user's physical activity and provides just-in-time feedback when the user becomes sedentary for more than 50 min [108]. This type of real-time monitoring can be supported by the use of diverse sensors. Current off-the-shelf smartphones include various sensors, such as microphones, GPS, motion sensors, a compass, a light sensor, and cameras. Sensor data can be passively collected in the background (known as opportunistic sensing; e.g., tracking a user's location traces), or users are asked to perform specific tasks for data collection (known as participatory sensing; e.g., taking a photo for food journaling) [109]. In particular, sensor and interaction data from smartphones or smartwatches facilitate a fine-grained understanding of user contexts and the detection of various events of interest, such as activity tracking with motion sensors and social interaction tracking with audio sensing or call/SMS log tracking. Wearable devices such as smartwatches and activity trackers offer a similar level of sensing, but the key departure is their support for sensing physiological signals, such as heart rate and skin temperature, which are useful for detecting stress and emotions (see Cowley *et al.*'s review [110]). Consumer electronics offer only limited capabilities for data collection, whereas wearable sensors for research purposes (e.g., Empatica E4 and Shimmer3) provide APIs for accessing raw data; however, their cost is an order of magnitude greater.

IoT devices can be classified based on their functionality and embeddedness. The major functionalities of IoT devices in domestic and office environments include IoT control (e.g., hubs or voice assistants), actuation (e.g., lighting and switches), and sensors (e.g., motion, temperature, and air quality). Most IoT devices are standalone products so that users can wirelessly control the connected door locks and thermostats through the Internet. Recent smart appliances, such as air conditioners, refrigerators, and air purifiers, are equipped with smart features such as state monitoring (e.g., operating condition logging and environment states) and remote control (e.g., temperature control). Unlike human activity monitoring with standalone IoT devices, smart appliances naturally come with sensors. As human appliance interactions are naturally part of activities of daily living, they provide useful information for health behavior monitoring (e.g., cognitive decline tracking [111]). Furthermore, smart appliances can help closely monitor home environments such as room temperature and air quality states (e.g., CO_2, PM2.5, and VOC). IoT devices or IoT-enabled appliances are typically connected to central hubs for integrated control over the Internet. Voice assistants include Amazon Echo and Google Home, which provide natural language support for information activities (e.g., Q&A) and device control (e.g., turning off bulbs). The type of IoT device control offered by IoT hubs, such as SmartThings Hub, provides a novel means of enabling DTx; for example, IoT devices can offer context-aware intervention for personalized sleep education (e.g., automatically turning off lights) [112].

B. Data-Driven Models for Engagement and Adherence

The digital intervention and passive sensor data are used to measure *user engagement in digital intervention* (or software usage) and *user adherence to behavioral changes*. Compared to "software engagement," adherence to behavioral changes is relatively easy to measure by tracking the user behavior based on user interactions (e.g., whether the content is consumed?) and passive sensing (e.g., whether physical activity has happened?) or by asking users to self-report. Diverse data-driven models for user engagement are discussed next; reviewing such models provides insight into the types of data that can be collected for data-driven DTx analytics.

A simple approach to measuring user engagement (and behavioral adherence) involves collecting self-reported feedback (e.g., level of satisfaction) [113]. Self-reports can be submitted at the end of experiments using well-known user engagement and usability scales, such as the system usability scale [114], usefulness, satisfaction, and ease of use questionnaire (USE) [115], and the user engagement scale (UES) [116]. These metrics consider diverse dimensions; for example, UES considers attention, usability, aesthetics, endurability, novelty, and involvement, and USE considers usefulness, ease of use, ease of learning, and satisfaction. In the case of digital health apps, information quality, trust, and security are also important [117]. While users typically answer these kinds of survey questionnaires at the end of an intervention period, it is also possible for participants to report their experiences on a shorter time scale (e.g., whenever some events happen or once a day) via EMA. An in-depth understanding of user engagement is difficult to acquire from surveys; in this case,
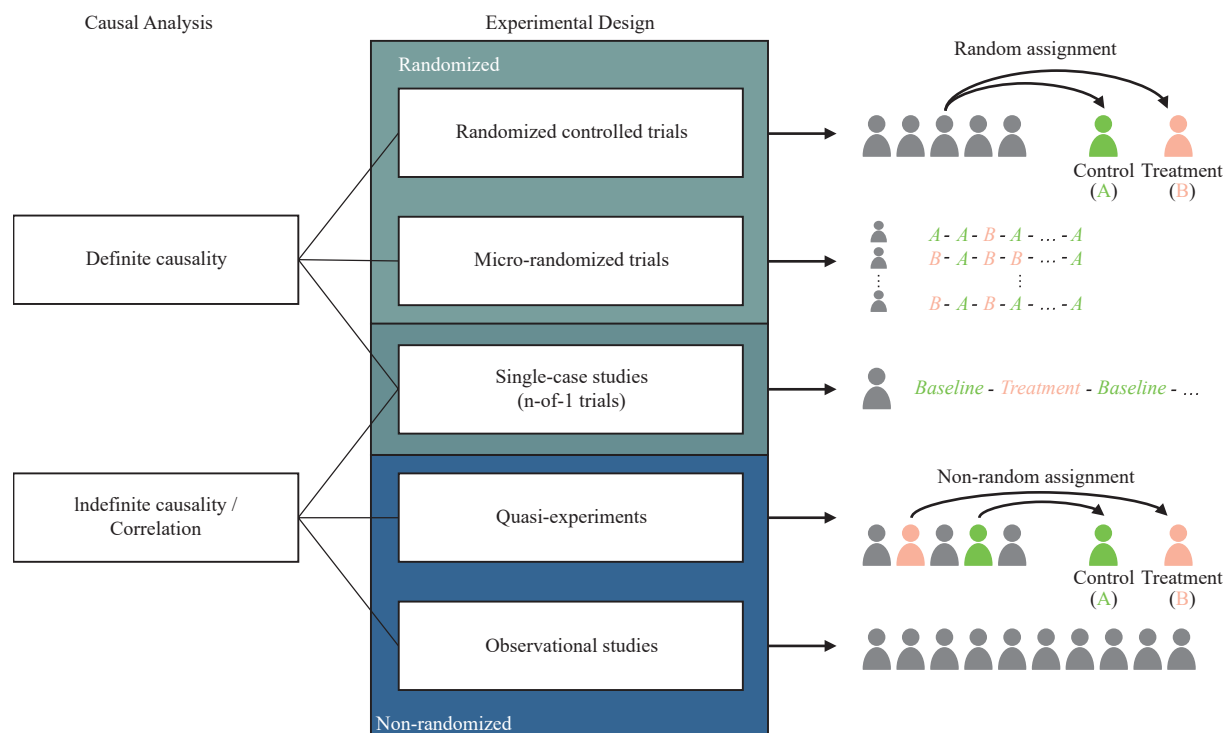
Fig. 5.   Experimental design and data analysis framework of DTx.

follow-up interviews and thematic analysis of the interview data are required [118].

Prior studies examined various usage metrics of digital interventions [80], [113], [119]. Basic usage metrics include the frequency of app use, user responses to proactive prompts, and on-demand usage that can be aggregated over a specific time window (e.g., per day or per week) [119]. In addition, interaction-level metrics can be considered. For a given use instance, the frequency of fine-grained interaction types, such as glancing, checking, and brief/detailed reviewing and their durations [80] can be summarized. In-depth interaction patterns can also be analyzed, such as the inter-checking intervals and receptivity to push notifications [113]. It is also possible to build a composite engagement index, a simple approach for a linear weighted sum of usage metrics [113] such as the number of pages viewed per session, frequency of app access, receptivity to push notifications, inter-checking intervals, and self-reported feedback (satisfaction).

In addition to such usage metrics, prior studies have examined the use of passive sensor data. Traditional engagement models based on flow theory mainly consider user allocation of attentional resources and their affective states. It is possible to measure a user's attention via behavioral and physiological sensor data, such as eye-tracking, facial expression, mouse movements, electrodermal activity (EDA), ECG, and electroencephalogram (EEG) data. Martey et al. [120] analyzed how passive sensor data (e.g., EDA and mouse movements/clicks) can be used to estimate self-reported engagement. Lascio et al. [121] explored how EDA data can be used to model a student's emotional engagement (i.e., enthusiasm and enjoyment) with the support vector machine (SVM). Mathur et al. [122] used EEG data to model user engagement from mobile context data and identified a strong correlation

between automatically detected engagement scores and users' subjective perception of engagement. Pielot et al. [123] considered consumption of recommended content as "engagement" and used machine learning (i.e., XGBoost) to identify latent engagement factors, which are based on diverse passive data collected ranging from mobile app usage to mobile sensing data. This type of mobile phone data can also be utilized on a larger scale (e.g., understanding people's behaviors and activities at the community level) [124]. Recent advances in mobile, wearable, and IoT technologies have enabled a wide array of novel health care consumer applications using big data analytics [125].

C. Experimental Design and Data Analysis Methods for DTx

Similarly to the development procedure of conventional therapeutics, the development and application of DTx depend heavily on the choice of experimental design and subsequent data analysis (Fig. 5). Various existing experimental designs suggested for eHealth or mHealth experiments can be applied to DTx to establish a causal relationship (i.e., evidence-based treatment). RCTs are the gold standard, whereby each experimental unit (or human subject) is randomly assigned to treatments (or conditions), and the outcomes are measured. The causal effect of the introduced treatment on the outcome variables is investigated by comparing the values of the outcome variables before and after the intervention while accounting for confounders.

The experimental design of data-driven DTx analytics further involves continuous passive data collection. Beyond establishing a causal relationship between the treatment and outcome variables, passively observed sensor data provide additional opportunities for DTx optimization by extracting various context and intervention variables (Fig. 6). Passive
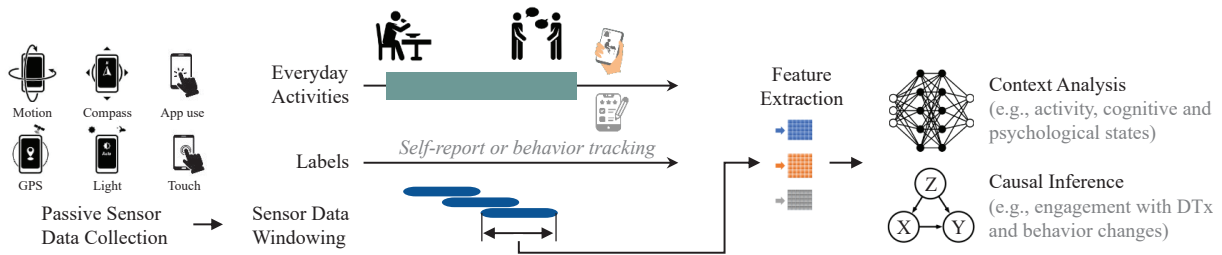
Fig. 6.   Data analysis pipeline for DTx analytics.

sensor data analyses can be used to understand intervention contexts, improve delivery mechanisms, evaluate intervention components, and personalize treatments (see Section V-E). In reality, RCTs are very expensive as they require hundreds (or even thousands) of participants, and are even less practical when DTx software has already been deployed. In the following subsection, both randomized and non-randomized experimental designs are reviewed along with the associated context analysis and causal effect inference methods applicable to data-driven DTx analytics.

1) Experimental Design for Data-Driven DTx Analytics: As with conventional therapeutics development, RCTs are the most appropriate experimental designs for DTx. In RCTs, the subjects are randomly assigned to either the experimental or control group. The treatment outcomes of the experimental group are compared with those of the control group to assess treatment effectiveness. In the RCT for DTx, the experimental group is given a digital or mobile platform, such as a smartphone application, and receives the desired intervention through the platform. In contrast, the control group can be defined in multiple ways: subjects in the control group may be given either conventional treatments, applications without any intended treatment effects, or no intervention at all [126]. Comprehensive reviews on RCTs for DTx regarding diseases such as depression, anxiety, and type 2 diabetes have been published [126]–[128].

Another type of randomized design applicable to DTx is the micro-randomized trial (MRT) [129], [130], which is a design proposed specifically for JITAI mHealth experiments [74]. In MRTs, the intervention a subject receives is randomized at every decision point, that is, a point in time during the experiment where an intervention may be effective. As a result, the interventions for each subject are sequentially randomized hundreds to thousands of times during the experiment, making the experiment resemble a sequential factorial design. Repeated randomizations for every subject point to two main advantages of using MRTs in DTx experiments. First, the time-varying causal effects of treatments can be assessed for the outcomes of interest. Second, both between-subject and within-subject comparisons are allowed in MRTs, making the experiments efficient. Randomizations may be stratified so that the number of interventions a subject receives is sufficient for all conditions (e.g., the stress level of subjects) [131].

In cases where an RCT is not applicable, designs such as quasi-experimental studies and observational studies, in which the assignment to the treatment or control group is not randomized, may be conducted for DTx. Quasi-experimental studies are experimental studies in which interventions are manipulated by the researchers without random assignments or full control over all extraneous confounding variables. Observational studies are non-experimental studies in which interventions are not controlled by researchers. Although these studies cannot draw any definite conclusions about the causal treatment effect because of the presence of confounders, they can be used to initially investigate the introduced interventions. For example, a mobile application to aid the treatment of diabetes may be provided to patients who would use the app freely without any explicit intervention by the researchers during the field trial [132]. Quasi-experimental studies such as pretest-posttest designs [133], [134] or matching designs [135], [136], have also been conducted to investigate the efficacy of mobile applications.

Single-case (also known as n-of-1) designs are a family of experimental designs applicable to DTx that contain aspects of both randomized and non-randomized trials. Single-case trials are experiments in which each subject acts as its own control, usually because the available sample size is very small. Similarly to MRTs, single-case trials capture the temporal dynamics within the study, as is often required by DTx studies. In general, single-case trials consist of two phases: the baseline and treatment. First, data are collected during the baseline period with no intervention, which serves as the control for the subject. Afterward, the intervention of interest is given to the subjects, and data for the same variables are collected. The causal effect of the intervention is estimated by comparing the treatment and baseline periods. In many cases, the baseline and treatment periods are repeatedly alternated (possibly with washout periods) for the "replication" of experiments or a clearer distinction of the causal effect of the treatment, by controlling for the confounding variables. In addition, different types of interventions or gradual changes in the intervention may be introduced after each baseline period, if necessary. When the baseline and treatment periods are randomly allocated within subjects, the trial is called an n-of-1 RCT. A more detailed review of single-case studies can be found in [137], [138].

2) Context Analysis Methods for DTx Analytics: The goal of context analysis is to use observational data to support exploratory tasks of investigating intervention contexts and behavioral routines or patterns, which can be captured using a set of observable characteristics of an individual through mobile and wearable devices such as activity trackers and smartphone loggers [139], [140]. The data sources range from passive sensor data (e.g., self-trackers and smartphone log-

ging) and social media use to active self-reporting (e.g., mood and stress).

Mobile and wearable devices allow the performing of continuous and unobtrusive measurements of user contexts and help to infer users' health and behaviors [141]. Prior studies have proposed several methods for analyzing users' intervention contexts and behavioral patterns. Here, "context" typically means "any information that can be used to characterize the situation of a person" [142]. A context refers to a situation and environment in which a device or user is situated by a set of relevant features, such as human factors (e.g., user, social setting, and tasks) and physical factors (e.g., location and infrastructure) [143]. Harari *et al*. [140] proposed a three-dimensional context model, which included social interaction, daily activities, and mobility patterns. Behavioral patterns in each dimension are further defined based on the data processing of user contexts (e.g., the duration of social interaction). Similarly, Mohr *et al*. [144] proposed a hierarchical contextual feature model in which low-level sensor data were transformed into low-level features that constituted high-level behavioral patterns. Most low-level features, such as location and activity types, are human-interpretable. High-level behavioral patterns may include intervention contexts or behavioral patterns related to mobility patterns (e.g., number of significant places visited) or interaction patterns (e.g., daily phone use frequency).

For context sensing, motion sensors can be used to detect various types of physical activity, such as movement [145], sleeping [146], eating [147], and agitation [148]. For example, a user's sedentary state can be easily calculated by comparing the arithmetic difference between accelerometer samples. Eating gestures can be recognized by applying machine learning (e.g., random forest) to wrist motion data measured using smartwatches [147]. GPS location traces can be processed to detect significant places (e.g., home and work) using clustering techniques [149]. In addition to physical activity tracking via sensors, prior studies used passive sensor data to infer cognitive and psychological states, such as depression [149] and social anxiety [150]. To identify the depressive symptoms of an individual, for example, mobility features (e.g., significant places visited) were extracted from smartphone GPS traces and personalized machine learning models with a support vector machine (SVM) classifier were tested [149]. One way of optimizing intervention contexts is to provide timely delivery of intervention with mobile, wearable, and IoT technologies, owing to obtrusive modality usage (e.g., visual, vibration, and sound notifications). However, interruptive messages result in productivity loss, increased stress, and time pressure [151], implying that less opportune delivery of intervention may lead to a low level of intervention adherence [152]. A user's behavioral routines can be leveraged to find opportune moments (e.g., activity transition times) [153]. A user's contextual model based on temporal and location contexts can be further used to define the complex rules for delivery timing.

3) Causal Inference Methods for Data-Driven DTx Analytics: Data-driven DTx analytics focuses on estimating the treatment effect under the potential outcomes (Neyman-Rubin causal) framework [154], [155], whereby the potential out-

comes (both factual and counterfactual) or a contrast (e.g., difference and ratio) are estimated either nonparametrically or by using models such as generalized estimating equation (GEE) and multilevel model (MLM). These models assume certain parametric regression form, in which the fitted model explains the causal effects of the treatment variables. For example, in MRTs, the average treatment effect under the potential outcome framework can be estimated through standard GEE or MLM regression, where the estimated coefficient of the treatment variable represents the mean difference in proximal outcomes between when the intervention is given and when it is not [129]. Furthermore, various estimation methods tailored for MRTs, such as the centered and weighted least-squares method [156], have also been proposed as a direct application of the models, which may lead to unstable estimation due to time-varying or endogenous covariates [131], [156], [157], [158]. Similarly, the GEE and MLM can be used to estimate the treatment effect in single-case studies with appropriate modifications [159], [160].

In the case of quasi-experimental or observational studies, it is difficult to identify definite causality because confounders are not controlled. Despite this limitation, it is common to compare the treatment and control groups to obtain a sense of the efficacy of an intervention [133], [134]. For example, as a pretest-posttest regarding the effect of a digital intervention on cardiovascular risk factors in postmenopausal women with obesity, the outcome variable (e.g., blood pressure) was measured before and after the experiment for both the treatment and control groups, and the difference in outcomes between the groups was compared [133]. A statistically significant difference between the groups suggests that the interventions are effective, although conclusions regarding the causal effect of the intervention cannot be made. Correlational analyses can also be conducted for the same purpose.

One possible approach for estimating causal effects using observational data is to adjust for the effects of confounders using methods such as stratification, matching, standardization, and weighting [161]. Adjustment methods can be applied either nonparametrically by calculating the necessary estimates from the data or parametrically by fitting some models to obtain the estimates (e.g., parametric G-formula and marginal structural models). It is common to use propensity scores [162] (or other types of balancing scores) to estimate the probability that the subject is placed in the treatment group, given covariate values, instead of directly using the observed variables. These methods attempt to create a pseudo-population or matched population from an observational dataset such that the treatment and control groups display a similar covariate distribution. For example, in a behavioral study using smartphone sensor data, the causal effects of various lifestyle factors (mobile phone use) on stress levels (emotional state) were analyzed by creating a matched population from context variables extracted from the sensor data [135], [136]. A practical tutorial on matching methods can be found in this reference [163].

Traditional adjustment methods are not generalizable to time-varying treatments and confounders, thereby resulting in biased estimates. To handle more complex longitudinal obser-

vational data with time-varying covariates, a more general class of models called G-methods for time-varying treatments was developed by Robins *et al.* [164]. These models, under appropriate assumptions, provide consistent estimates on contrasts of average treatment effects in scenarios with time-varying treatments and confounders by taking into consideration the "history" (the value of treatment and confounding variables up to the time point of interest) of subjects.

This review focuses on the potential outcomes framework. In contrast, the causal inference developed by Pearl [165], based on causal graphs and structural causal models, can also be implemented for similar purposes. In Pearl's framework, each node represents a variable, and the directed edges represent direct causation from one variable to another. The causal relationships encoded in *causal graphs* can be translated into structural models and various causal inference tasks, such as causal structure discovery, which retrieves the underlying causal graph from observational data, and then treatment effect estimation from observational data can be conducted.

Another notion of causality often encountered in the literature is *predictive causality*. Granger causality [166], one of the most common concepts of predictive causality, measures whether a time series is predictive of another by testing whether the lagged values of one variable have predictive power for the upcoming values. In the case of dynamic systems where variables exhibit deterministic relationships, convergent cross mapping [167] can be used to discover the causality of two time series by leveraging time-delayed embeddings. These methods have been used to explore predictive causality in mobile data [168], [169]. Readers can find a comparative analysis of these approaches in [170].

D. Data Visualization for DTx

Data visualization aims to visually represent datasets to help people perform exploratory tasks more effectively [171]. Existing data visualization tools often support various user interactions such as changing visual encodings and navigating datasets (e.g., zooming, slicing, and projecting), thereby facilitating data-driven reasoning or decision-making, which is referred to as visual analytics [172].

Data-driven DTx analytics is based on personal big data collected from mobile and wearable devices. Passively collected DTx datasets are characterized by a high volume and velocity data-stream of sensor and interaction data with different spatiotemporal granularities (e.g., temporal: event time or day/week/month windowing, spatial: GPS points, places, or regions), semantic hierarchies (e.g., semantic grouping of sensors), and data types (quantitative, categorical, and unstructured). Visual analytics tools should support visual data exploration (e.g., charting datasets) and automated data analysis (e.g., data mining and machine learning) [172]. Visual data exploration, on the other hand, helps users check data quality and test research hypotheses visually. In addition, insightful information can be extracted by interacting with the data (e.g., examining and comparing different datasets and, zooming in on different parts). These insights, in turn, help build better models for data analyses and visualization explorations. In data-driven DTx analytics, this type of value-sensitive visual-

ization is critical. Clinicians or developers may want to inspect user contexts, DTx usage, and behaviors not only to understand user engagement and behavioral adherence but also to acquire design insights for DTx improvements.

When visualizing time-oriented data, the first step is to characterize the key characteristics of time-series data, such as temporal primitives (i.e., time points vs. intervals) and the structure of time (e.g., linear vs. cyclic as in seasons of the year) [173]. As reviewed by Brehmer *et al.* [174], effective timeline visualization requires careful exploration of the design space such as visual representation (e.g., linear, radial, grid, or spiral), scale of a timeline (e.g., sequential, chronological, relative, or logarithmic), and layout of views (e.g., single vs. multiple timelines, and segmented timelines). In addition, analytic reasoning for time-series data can be facilitated by supporting diverse user-interaction techniques (e.g., zooming, faceting, focus, and context) [175]. Thus, data visualization support for DTx must carefully consider data characteristics of mobile and wearable data and analytic reasoning tasks.

Visualizing a large volume of personal big data is challenging; thus, it is important to support visual summaries of multidimensional time-series data. A simple visualization of the data streams collected from mobile and wearable devices would not be scalable; for example, a minute-level "behavior-gram" of smartphone data [176] could be quickly cluttered as the number of data streams increases. For effective visualization, DataMD [177] provides a visual dashboard for clinicians with visual summaries and trend lines of user activities for data-driven consultation. Interactive visualization allows users to compare multiple relevant timelines to help explore rich contextual data (e.g., identifying stressors by examining stress levels across different contexts over time [178]). Advanced machine learning techniques can be used to find key relevant features such as behavioral markers related to cognitive decline [176]. Event sequence mining can be used to uncover temporal patterns such as the event patterns before the occurrence of smoking [179].

E. DTx Design and Development Process

Digital therapeutics can be developed and improved during the product lifecycle. In this section, a review of DTx design and development processes is provided to illustrate how data-driven DTx analytics can help improve DTx services (Fig. 7).

1) DTx Design and Development Process Review: The concept of DTx is highly related to that of digital health and persuasive technology in that software is used to persuade users to change their behaviors. Thus, a literature review is provided regarding the design and development process: a) plan and target setting, b) design and development, and c) evaluation of efficacy.

Plan and target setting: The first step is to identify the target medical disorder or disease. Fogg mentioned "a simple target behavior to change" as the first consideration for designing a successful persuasive technology [81]. Since the design process of DTx is mainly rooted in pervasive technology, a target disease or behavior should be selected first. The behavioral changes mentioned here could be the detailed goals of a larger objective (i.e., dealing with medical disorders). For
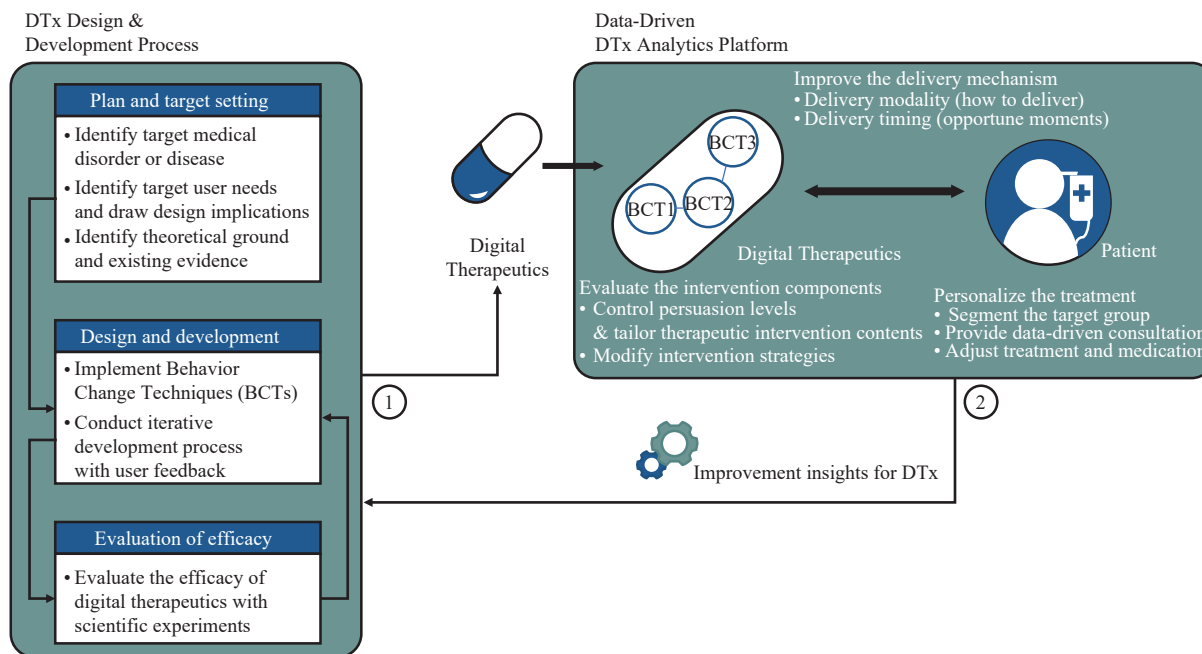
Fig. 7.   DTx Design and Development Process Overview: (1) DTx design and development process with a) Plan and target setting, b) Design and development, c) Evaluation of efficacy, and (2) Data-driven DTx analytics for acquiring insights for improvement in three ways—a) Improve the delivery mechanism, b) Evaluate the intervention components, and c) Personalize the treatment. These insights can be re-fed back into the development process to iteratively improve DTx.

example, patients with prediabetes can set detailed goals for exercise or diet management to prevent diabetes. After setting the target disease, researchers should identify the needs of the target user and draw design implications. It is crucial to understand the patients' point of view on digital health product use to implement what they need for the behavior change through "patient-centered design" methods, which involve patients in the process of development [180]. In addition, it is necessary to identify the theoretical basis and existing evidence underlying the design of the intervention, which provides the rationale for selecting specific target diseases and behaviors, target users, and behavioral change techniques [19]. Evidence from existing research can be collected using a systematic review of RCTs, or by testing relevant proven methods before development [181].

Design and development: The second step begins with the implementation of behavior change techniques. The behavior change technique (BCT) refers to "an observable, replicable, and irreducible component of an intervention designed to alter or redirect causal processes that regulate behavior [182]" For example, for the purpose of physical activities, various BCTs, such as providing guidance for a specific behavior, providing feedback on the user's performance, and setting goals of behavior and outcome are widely used [183]. During this process, the number and type of BCTs that can be applied to DTx should be determined based on the identified target diseases, users, and theoretical/evidence bases. BCTs are then instantiated with certain intervention components to allow patients to interact directly with them. During this process, the medium to use, timing, and frequency of delivery of the intervention should also be considered [12]. As an iterative development process, pilot tests should be conducted on potential users (in

this case, patients with the target disease) to examine the acceptability and feasibility of DTx intervention, user engagement, and adherence to the instructions.

Evaluation of efficacy: The third step is to evaluate the efficacy of the DTx intervention. The assessment of the efficacy can measure different endpoints: 1) how much the patient behavior changed (i.e., proximal outcome) and 2) the degree of disease prevention or treatment (i.e., distal outcome). Randomized controlled (or clinical) trials are favored for assessing the relative effectiveness of DTx between treatment and control groups to reduce potential biases (e.g., selection bias). According to a report, RCTs and meta-analysis studies evaluating the clinical efficacy of digital health apps are gradually increasing [39].

2) Data-Driven DTx Analytics for Optimizing DTx Delivery: Data-driven DTx analytics can improve the efficacy of existing DTx by uncovering insights to improve therapeutics and apply them to the iterative development cycle based on the patient data. Since existing clinical trials have evaluated the efficacy of digital health interventions mainly with the difference in endpoints, it becomes a challenge to understand how certain elements of DTx induce differences and which factors beyond DTx affect patients' health concurrently [23]. Beyond efficacy evaluation, data-driven DTx analytics can play an important role in further improving DTx by a) improving the delivery mechanism, b) evaluating the intervention components, and c) personalizing treatments.

Improve delivery mechanisms: For the delivery of health interventions, it is necessary to determine the channels of interaction between DTx and the patient. The appropriate channels of sensory input and output, or modalities, may vary depending on user preference (i.e., the preferred device or

modality) or context (i.e., whether the intervention is palpable or interruption by sound notification is allowed) [13]. The patient's context and the actual response for each modality can be analyzed in real time to suggest an effective method that most likely changes the patient's behavior. Modality selection can be automatic, considering the trade-off between the most effective modality and its availability (e.g., converting sound notification into mute and LED light during a meeting). Depending on the context, the patient may or may not be aware of the DTx instructions and hence may not adhere to them. Therefore, it is crucial to detect "opportune moments" to interrupt patients to deliver timely health interventions to maximize efficacy. These moments can be determined by context-sensing methods specific to the platform or feature extraction by machine learning [13]. During this process, sensor data may be utilized from smartphones (e.g., ambient light, sound, or acceleration of the device) for interaction (e.g., touching or sliding the screen) to understand the context of the user, as suggested in [184]. It is also possible to divide the delivery process of the intervention into several stages to identify the point at which the process fails [15]. As suggested in this study, whether the patient a) perceives the intervention, b) is available to respond, c) adheres to the intervention, and d) performs the target behavior can be affected by several contexts. Thus, contextual data are collected and utilized to build a model that can improve delivery mechanisms.

Evaluate the intervention components: During the development process of DTx, BCTs are determined based on a theoretical basis or existing proven research. However, such BCTs may not result in the same efficacy because of differences in the target medical disorder, target user traits, or the context of therapeutic use. Data-driven DTx analytics can be used to evaluate the extent to which each BCT component affects the target behavior (proximal outcome). A component-wise evaluation is less practical because of its complexity and cost. Instead, the observed sensor data can be used to determine this type of component-wise efficacy via a pseudo-experiment design. This helps determine the relative importance of components and guides designers and clinicians on how to improve the components or to better evaluate BCTs (experimental design). In addition, it may be possible to identify inter-relationships and causal relationships of the components, similarly to a model describing correlations and interactions among health-related attitudes [185]. This evaluation can help develop a new strategy, such as emphasizing certain BCT components by similar patient groups or deploying motivators to enhance their use. Compliance with persuasion for DTx may vary depending on the patient's personality traits, such as extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience [186]. The data collected can be used to create a "persuasion profile" to predict how many different types of persuasion principles (e.g., authority, consensus, commitment, or scarcity) would affect patients with different personalities [187]. Thus, the persuasion level can be controlled by changing the intensity of the wording, the frequency of sending persuasive messages, and personalizing persuasion principles to each individual. In addition, the patterns of interaction can differ with intervention content

because the patient may have difficulty in understanding the content, lose interest in the content, or not be motivated to learn. Data-driven analytics then utilizes interaction patterns to design intervention coursework or recommend content that may engage the patient's interest. Recently, Yue *et al.* [188] reviewed widely used techniques for content recommendation, such as content-based (i.e., content with similar features), collaborative filtering-based (i.e., modeling user behavior), and hybrid methods. These approaches can also be adapted to the intervention content recommendations in DTx.

Personalize the treatment: DTx can be further improved or "personalized" through an understanding of the target group via contextual analytics. Patients can be divided into subgroups based on their engagement in or adherence to therapeutics. For example, Alshurafa *et al.* [80] classified clusters of patients who showed similar health behavior changes in four categories (i.e., "healthy and steady," "unhealthy and steady," "decliners," and "improvers"). This approach may help predict changes in health activities by considering the characteristics of patient groups to formulate appropriate strategies. In addition, more personalized treatments can be achieved by adding personal traits. For instance, the outcome of behavior change can be analyzed in relation to certain variables, such as demographics, health status, and psychosocial features [189]. This process enables the analysis of more specific target segments and even covers niche patient groups using upgraded DTx. The analytics platform can also help health coaches understand patient engagement and adherence to DTx based on collected data (e.g., the patients' context, therapeutic usage pattern, and proximal and distal outcomes). Moreover, the platform may suggest the advice needed for the patient by visualizing activity data. With this information, the coach can provide data-driven, personalized discussions to the patient and improve the efficacy of therapeutics. DTx can be used independently or along with other traditional treatments. If the patient takes medicine as part of DTx (e.g., Pear Therapeutics' substance use disorder intervention), the analytics can be used to increase the effectiveness of the medication in use and result in better outcomes. For example, in the case of depression, DTx allows the patient to record the degree of depression, quality of sleep, and activities performed. Based on these data, a clinician can control the type, combination, or dose of medication when the patient visits the hospital.

VI. RESEARCH DIRECTIONS FOR DATA-DRIVEN DTx ANALYTICS

A. Data Collection and Data-Driven Delivery Optimization

Data Quality Management: While conducting a large-scale experiment in the wild, a general-purpose data platform guarantees that all necessary data from users are collected with data quality support (e.g., format, completeness, and freshness [190]). When data quality is compromised, the analytics platform can notify key stakeholders (e.g., data providers, managers, or DTx developers) for further intervention (e.g., checking mobile apps or fixing software errors). Moreover, the analytics platform may need to detect a user's abnormal behavior for quality management (i.e., injecting fake data to

maximize rewards or violation of experimental guidelines). These abnormalities can be easily captured by outlier detection with suitable feature tracking (e.g., local outlier detection [191] and machine learning [192]). Existing CTMSs or data collection platforms can consider implementing additional features to enhance data quality management.

Intervention Delivery Optimization: Another research direction related to receptivity improvement of analytics platforms is opportune moment detection and modality selection; these analyze tailoring variables (i.e., user's behavioral and contextual data) and intervention components to find the most effective time that maximizes user engagement. This includes deferring the intervention until the predicted deadline [193]. Furthermore, the analytics platform requires selecting an appropriate modality in a multi-device environment (e.g., laptops, smartphones, smartwatches, and IoT appliances) to consider user preferences, device accessibility, and properties of the intervention [194]. Finally, finding an opportune moment and modality selection should be cross-optimized to improve receptivity to the DTx software, which is an open and interesting research area.

B. Causal Inference for Data-Driven DTx Analytics

In DTx scenarios, observational studies are conducted along with randomized studies because of the cost or difficulty of conducting complex experiments. For example, the number of conditions increases exponentially with the number of factors (e.g., intervention components). Recent advances in artificial intelligence and deep learning have led to high-performing causal inference models for observational data, which handle the lack of randomized controlled environments through model architecture [195]–[200]. In these studies, individual treatment effect estimation was performed through counterfactual inference, in which machine learning models, including deep neural networks [196], [197], [199], [200] and tree-based models [195], [198], were used to estimate all potential outcomes and thus the causal effect. Counterfactual inference models that use sequential models such as recurrent neural networks have also been proposed to handle the bias from time-varying confounders [201]–[203]. Similar counterfactual inference frameworks can be applied to data-driven DTx analytics in which the effect of treatments delivered through digital or mobile platforms is estimated through deep networks based on passively collected DTx datasets. Various definitions of treatments can be explored and implemented accordingly, depending on the causal analysis of interest, through effect estimation models. For example, treatment variables can be defined based on the engagement levels of the DTx software, or various sets of actions can be observed via passively collected data. Regardless of how the treatment variables are defined, an appropriate model architecture that incorporates all aspects of passively collected DTx data is required to estimate all potential outcomes for accurate causal analysis of DTx.

C. Data Visualization for Data-Driven DTx Analytics

One of the critical challenges in data visualization is the volume of personal stream big data. During a clinical trial, a stream of sensor and interaction data can be collected from hundreds or thousands of people over a few months. As discussed in Section VI-D, it is also possible to collect data after deployments for continuous improvements (from those who opt-in for data collection). Prior studies explored various interactive visualization techniques to efficiently navigate through large-scale disease or patient datasets [204], [205]; for example, CareFlow helps visualize the outcome data of 50 000 patients in a tree-like timeline and provides a high-level overview of similar groups of patients [205]. Large-scale datasets are processed to extract hundreds or thousands of features for contextual and causal analytics. Visualization of such high-dimensional data is challenging. As in existing machine learning studies, dimensionality reduction techniques such as principal component analysis or t-distributed stochastic neighbor embedding can be used for dataset exploration and annotation. Leveraging domain knowledge as well as prior data-driven DTx analytics experience can also help focus on specific sensor and interaction data types. As illustrated earlier, DTx design and development are mainly led by prior domain knowledge on theoretical grounds and existing evidence [19]. According to a recent survey on depression, several key behavioral features include location, physical activity, and sleep [206]. Data-driven DTx analytics, including correlation analyses and causal inference, allows for further reduction of the feature space. As a result, clinicians can manually examine the reduced feature space to effectively explore hypotheses on contextual and intervention factors affecting engagement and behavioral adherence. It is envisioned that artificial intelligence-inspired ranking mechanisms will automatically generate a ranked list of hypotheses to be examined.

D. DTx Design and Development Process

The existing literature on DTx design mostly focuses on design principles, and there is a lack of practical guidelines on how software development and evaluation methods can be incorporated into the design and development process. As shown earlier, the US FDA introduced a pre-certification pilot program that certifies SaMD "developers" instead of SaMD "products" to manage the total product lifecycle (TPLC), which "enables the evaluation and monitoring of a software product from pre-market development to post-market performance, along with the continued demonstration of the organization's excellence" [61]. It is important to collect and analyze real-world datasets for DTx to continuously improve product safety and effectiveness and to deal with potential risks.

The challenge is to establish user-centered agile software development (UCASD) practices and principles for DTx design, development, and evaluation, as recommended by Pach *et al.* [207] (e.g., minimizing the complexity of DTx design and jointly conducting app development and evaluation). Well-known UCASD principles can be applied to DTx. Software design and development occur in short iterations, with incremental improvements. Design, development, and evaluation occur in parallel interwoven tracks, and stakeholders (e.g., patients, clinicians, and developers) are actively involved in the entire process [118], [208]. In UCASD, evalu-

ation can occur at various stages: initial (pre)design phases (e.g., plan/target setting and implementation of behavior change techniques), iterative development of DTx software, clinical evidence evaluation (for clinical trials), and in-the-wild deployments. Existing user-centered design techniques (e.g., small-scale user studies by interviewing participants) can consider "data-driven DTx analytics." It is important to systematically examine how diverse datasets collected during different evaluation stages can be used by key stakeholders to acquire design insights for DTx improvements (e.g., data mining for usability issue identification [209]). For example, during the clinical evaluation stage, RCTs can be conducted using passive data collection for data-driven analytics. When considering the TPLC of DTx, data-driven analytics helps a company not only with continuous monitoring of product safety and risk factors but also in providing opportunities for iterative design improvements. Further case studies on DTx design and development processes using data-driven DTx analytics should be conducted to establish practical DTx design and development practices and principles.

The review and discussion in this work mostly focused on conventional mobile and wearable application designs for DTx. It is increasingly important to consider new computing platforms for BCT delivery, such as virtual reality and metaverse (e.g., Facebook's Horizon Workrooms, Microsoft Mesh, and NVIDIA Omniverse) [210] and conversational agents (e.g., Google Assistant) [211]. Further studies must be conducted to understand user engagement and adherence patterns in novel platforms, such as tracking an avatar's behaviors in the metaverse for data-driven DTx analytics. In addition, the delivery of therapeutic interventions can be extended using physically embodied smart agents, such as robots. In terms of human-robot interaction, diverse responses are collected from users, such as visual/vocal nonverbal cues or subjective feelings during the interaction, which are then utilized to evaluate and improve interventions [212], [213].

E. Privacy and Ethics in DTx Analytics

Given the DTx analytics uses a vast amount of personal data collected from mobile, wearable, and IoT devices (e.g., GPS locations, phone usage, and emotion samples), privacy and ethics must be carefully evaluated. Recent studies examined user motivations and concerns when collecting data on mobile and wearable devices [214], [215]. The two main reasons why users provide their data are financial gain and altruistic benefits (e.g., scientific advances) [215]. According to Rooksby *et al.* [214], the major concerns are related to negative user experiences of data collection; e.g., privacy concerns of sensitive data collection, negative thoughts and feelings about self-reporting, and detrimental effects on device performance. Although the majority of data contributors are not typically concerned with privacy, Lee *et al.* [215] further showed that privacy concerns are related to feelings of surveillance, the identification of daily routines, and data breach.

For ethical reasons, Huh-Yoo *et al.* [216] stressed the significance of communicating with the study participants as part of the informed consent process by elaborating on the known potential risks (as well as their additional concerns) and how

to mitigate those risks. In DTx settings, it is critical to communicate the types of behavioral and sensor data that are gathered, whether any of it contains any personally identifiable information, and the potential risks associated with sharing sensor data. However, it was found that users frequently have incorrect mental models regarding the types of data collected from mobile and wearable devices, and how they are used to build AI models [215]. Therefore, researchers must provide intuitive explanations about the types of data collected and allow users to access their own data by supporting interactive data visualization tools (e.g., minute-level behaviorgram of sensor data streams [217]). In fact, DTx analytics could leverage the MyData vision that empowers users by offering *direct* access and control of their own data [218]. The MyData vision extends the EU's General Data Protection Regulation (GDPR) that intends to enhance users' control and rights over personal data. Further study is required to develop novel user-friendly tools that allow users to directly access and control their own data as part of DTx data analytics platforms, similar to Privacy Bird [219], a user-friendly web data privacy management tool for the Platform for Privacy Preferences (P3P).

Data collection requires informed consents from participants where researchers and clinicians must inform participants about the risks and benefits of study participation. Before collecting any data, one-time informed consents are normally obtained. Recent research has supported the usage of dynamic informed consents, which enable users to adaptively modify their consents in response to their contextual needs, such as turning off GPS while visiting hospitals [220]. Future research should investigate the design space of dynamic consents in mobile and wearable environments, much like earlier studies that thoroughly studied the design space of privacy notice and choices [221], [222]. We anticipate that implementing dynamic consents will enhance the MyData vision in DTx contexts.

VII. Conclusion

Digital therapeutics as SaMD aims to cure diseases and improve health conditions, which is a major departure from existing wellness products. Health interventions are delivered through digital technologies (e.g., mobile content, chatbots, and push notifications); thus, it is critical to analyze and optimize the engagement and receptivity to DTx delivery systems. We proposed a data-driven DTx analytics framework that allows researchers and practitioners to collect mobile, wearable, and IoT data as part of field experiments and to perform contextual analytics and causal inference for DTx delivery optimization. We reviewed the core components of data-driven DTx analytics such as data collection and user management, data-driven modeling for DTx usage and behavioral adherence, experimental design and contextual/causal analytics, data visualization, and DTx design optimization. Finally, we discussed research directions for data-driven DTx analytics, such as causal inference, data visualization, delivery optimization, design process optimization, and the privacy and ethics of data-driven DTx. Over the last few decades, there have been significant advances in mobile sensing and machine learning, and observational data have provided new opportuni-

ties for DTx development and optimization. However, data-driven DTx analytics faces novel challenges such as privacy risks due to extensive data surveillance and development and operation risks due to the heterogeneity of mobile platforms and IoT devices (e.g., iOS vs. Android). Despite these challenges, we hope that our work serves as a foundation for this new research direction, and we call for further research on data-driven DTx analytics.

REFERENCES

[1]   O. Sverdlov, J. Van Dam, K. Hannesdottir, and T. Thornton-Wells, "Digital therapeutics: An integral component of digital innovation in drug development," *Clin. Pharmacol. Ther.*, vol. 104, no. 1, pp. 72–80, Jul. 2018.

[2]   Digital Therapeutics Alliance, "Digital therapeutics: Combining technology and evidence-based medicine to transform personalized patient care," 2018. [Online]. Available: https://dtxalliance.org/wp-content/uploads/2021/01/DTA_DTx-Definition-and-Core-Principles.pdf, Accessed on: Mar. 6, 2022.

[3]   C. C. Quinn, M. D. Shardell, M. L. Terrin, E. A. Barr, S. H. Ballew, and A. L. Gruber-Baldini, "Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control," *Diabetes Care*, vol. 34, no. 9, pp. 1934–1942, Sep. 2011.

[4]   A. N. C. Campbell, E. V. Nunes, A. G. Matthews, M. Stitzer, G. M. Miele, D. Polsky, E. Turrigiano, S. Walters, E. A. McClure, T. L. Kyle, A. Wahle, V. Veldhuisen, B. Goldman, D. Babcock, Q. Stabile, T. Winhusen, and U. E. Ghitza, "Internet-delivered treatment for substance abuse: A multisite randomized controlled trial," *Am. J. Psychiatry*, vol. 171, no. 6, pp. 683–690, Jun. 2014.

[5]   S. C. Sepah, L. H. Jiang, R. J. Ellis, K. McDermott, and A. L. Peters, "Engagement and outcomes in a digital diabetes prevention program: 3-year update," *BMJ Open Diabetes Res. Care*, vol. 5, no. 1, p. e000422, Sep. 2017.

[6]   S. H. Kollins, D. J. DeLoss, E. Cañadas, J. Lutz, R. L. Findling, R. S. E. Keefe, J. N. Epstein, A. J. Cutler, and S. V. Faraone, "A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): A randomised controlled trial," *Lancet Digit. Health*, vol. 2, no. 4, pp. e168–e178, Apr. 2020.

[7]   Grand View Research, "Digital therapeutics market size, share & trends analysis report by application (diabetes, obesity, CVD), by end user (patients, providers, payers, employers), by region, and segment forecasts, 2022‒2030," 2022. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/digital-therapeutics-market, Accessed on: Mar. 6, 2022.

[8]   D. A. Greenwood, M. Gee, K. J. Fatkin, and M. Peeples, "A systematic review of reviews evaluating technology-enabled diabetes self-management education and support," *J. Diabetes Sci. Technol.*, vol. 11, no. 5, pp. 1015–1027, Sep. 2017.

[9]   R. J. Widmer, N. M. Collins, C. S. Collins, C. West, L. O. Lerman, and A. Lerman, "Digital health interventions for the prevention of cardiovascular disease: A systematic review and meta-analysis," *Mayo Clin. Proc.*, vol. 90, no. 4, pp. 469–480, Apr. 2015.

[10]   C. Hollis, C. J. Falconer, J. L. Martin, C. Whittington, S. Stockton, C. Glazebrook, and E. B. Davies, "Annual research review: Digital health interventions for children and young people with mental health problems—A systematic and meta—Review," *J. Child Psychol. Psychiatry*, vol. 58, no. 4, pp. 474–503, Apr. 2017.

[11]   D. Liu, F. Yang, F. Xiong, and N. Gu, "The smart drug delivery system and its clinical potential," *Theranostics*, vol. 6, no. 9, pp. 1306–1323, Jun. 2016.

[12]   D. C. Mohr, S. M. Schueller, E. Montague, M. N. Burns, and Rashidi, "The behavioral intervention technology model: An integrated conceptual and technological framework for eHealth and mHealth interventions," *J. Med. Internet Res.*, vol. 16, no. 6, p. e146, Jun. 2014.

[13]   U. Lee, K. M. Han, H. Cho, K. Chung, H. Hong, S. J. Lee, Y. Noh, S. Park, and J. M. Carroll, "Intelligent positive computing with mobile, wearable, and IoT devices: Literature review and research directions," *Ad Hoc Netw.*, vol. 83, pp. 8–24, Feb. 2019.

[14]   B. K. Wiederhold, "Data-driven digital therapeutics: The path forward," *Cyberpsychol. Behav. Soc. Netw.*, vol. 24, no. 10, pp. 631–632, Oct. 2021.

[15]   W. Choi, S. Park, D. Kim, Y. K. Lim, and U. Lee, "Multi-stage receptivity model for mobile just-in-time health intervention," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 2, p. 39, Jun. 2019.

[16]   J. Zhou, Y. H. Zhou, B. C. Wang, and J. Y. Zang, "Human-cyber-physical systems (HCPSs) in the context of new-generation intelligent manufacturing," *Engineering*, vol. 5, no. 4, pp. 624–636, Aug. 2019.

[17]   J. C. Kvedar, A. L. Fogel, E. Elenko, and D. Zohar, "Digital medicine's march on chronic disease," *Nat. Biotechnol.*, vol. 34, no. 3, pp. 239–246, Mar. 2016.

[18]   N. A. Patel and A. J. Butte, "Characteristics and challenges of the clinical pipeline of digital therapeutics," *NPJ Digit. Med.*, vol. 3, no. 1, p. 159, Dec. 2020.

[19]   T. Webb, J. Joseph, L. Yardley, and S. Michie, "Using the internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy," *J. Med. Internet Res.*, vol. 12, no. 1, p. e4, Feb. 2010.

[20]   S. C. Sepah, L. H. Jiang, and A. L. Peters, "Long-term outcomes of a web-based diabetes prevention program: 2-year results of a single-arm longitudinal study," *J. Med. Internet Res.*, vol. 17, no. 4, p. e92, Apr. 2015.

[21]   E. Capobianco, "On digital therapeutics," *Front. Digit. Humanit.*, vol. 2, p. 6, Nov. 2015.

[22]   C. Sepah, "The double-edged sword of digital health (the downside will surprise you)," 2017. [Online]. Available: https://medium.com/goactualize/the-double-edged-sword-of-digital-health-the-downside-will-surprise-you-8eaf5c3c84df, Accessed on: Mar. 6, 2022.

[23]   World Health Organization, "WHO guideline: Recommendations on digital interventions for health system strengthening," World Health Organization, Geneva, Switzerland, 2019.

[24]   U.S. Food & Drug Administration, "What is digital health?," 2020. [Online]. Available: https://www.fda.gov/medical-devices/digital-health-center-excellence/what-digital-health, Accessed on: Mar. 6, 2022.

[25]   Digital Medicine Society, "Defining digital medicine," 2021. [Online]. Available: https://www.dimesociety.org/about-us/defining-digital-medicine/, Accessed on: Mar. 6, 2022.

[26]   Digital Therapeutics Alliance, "Digital health industry categorization," 2019. [Online]. Available: https://dtxalliance.org/wp-content/uploads/2019/11/DTA_Digital-Industry-Categorization_Nov19.pdf, Accessed on: Mar. 6, 2022.

[27]   T. Aungst and J. Murdock, "What's the difference between digital health, digital medicine, and digital therapeutics?," 2021. [Online]. Available: https://www.goodrx.com/healthcare-access/telehealth/digital-therapeutics-vs-digital-medicine-vs-digital-health, Accessed on: Mar. 6, 2022.

[28]   G. E. Iyawa, M. Herselman, and A. Botha, "Digital health innovation ecosystems: From systematic literature review to conceptual framework," *Procedia Comput. Sci.*, vol. 100, pp. 244–252, Oct. 2016.

[29]   U.S. Food & Drug Administration, "Enforcement policy for non-invasive remote monitoring devices used to support patient monitoring during the coronavirus disease 2019 (COVID-19) public health emergency (revised)," 2020. [Online]. Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/enforcement-policy-non-invasive-remote-monitoring-devices-used-support-patient-monitoring-during, Accessed on: Mar. 6, 2022.

[30]   Digital Therapeutics Alliance, "Digital therapeutics: Reducing rural health inequalities," 2020. [Online]. Available: https://dtxalliance.org/wp-content/uploads/2021/01/DTA_Rural-Health_r13_110220.pdf, Accessed on: Mar. 6, 2022.

[31]   D. Morrison, S. Wyke, K. Agur, E. J. Cameron, R. I. Docking, A. M. MacKenzie, A. McConnachie, V. Raghuvir, N. C. Thomson, and F. S.

Mair, " Digital asthma self-management interventions: A systematic review," *J. Med. Internet Res.*, vol. 16, no. 2, p. e51, Feb. 2014.

[32] A. Khaylis, T. Yiaslas, J. Bergstrom, and C. Gore-Felton, "A review of efficacious technology-based weight-loss interventions: Five key components," *Telemed. J. E. Health*, vol. 16, no. 9, pp. 931–938, Nov. 2010.

[33] K. Ghorai, S. Akter, F. Khatun, and Ray, " mHealth for smoking cessation programs: A systematic review," *J. Pers. Med.*, vol. 4, no. 3, pp. 412–423, Jul. 2014.

[34] M. Aapro, Bossi, A. Dasari, L. Fallowfield, Gascón, M. Geller, K. Jordan, J. Kim, K. Martin, and S. Porzig, "Digital health for optimal supportive care in oncology: Benefits, limits, and future perspectives," *Support. Care Cancer*, vol. 28, no. 10, pp. 4589–4612, Oct. 2020.

[35] M. J. Choi, H. Kim, H. W. Nah, and D. W. Kang, "Digital therapeutics: Emerging new therapy for neurologic deficits after stroke," *J. Stroke*, vol. 21, no. 3, pp. 242–258, Sep. 2019.

[36] G. Abbadessa, F. Brigo, M. Clerico, S. De Mercanti, F. Trojsi, G. Tedeschi, S. Bonavita, and L. Lavorgna, " Digital therapeutics in neurology," *J. Neurol.*, vol. 269, no. 3, pp. 1209–1224, May 2022.

[37] J. M. Simoni, B. A. Kutner, and K. J. Horvath, "Opportunities and challenges of digital technology for HIV treatment and prevention," *Curr. HIV/AIDS Rep.*, vol. 12, no. 4, pp. 437–440, Sep. 2015.

[38] C. E. Canan, M. E. Waselewski, A. L. D. Waldman, G. Reynolds, T. E. Flickinger, W. F. Cohn, K. Ingersoll, and R. Dillingham, "Long term impact of *PositiveLinks*: Clinic-deployed mobile technology to improve engagement with HIV care," *PLoS One*, vol. 15, no. 1, p. e0226870, Jan. 2020.

[39] M. Aitken, B. Clancy, and D. Nass, " The growing value of digital health: Evidence and impact on human health and the healthcare system," 2017. [Online]. Available: https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health, Accessed on: Mar. 6, 2022.

[40] B. Meyer, T. Berger, F. Caspar, C. Beevers, G. Andersson, and M. Weiss, " Effectiveness of a novel integrative online treatment for depression (Deprexis): Randomized controlled trial," *J. Med. Internet Res.*, vol. 11, no. 2, p. e15, May 2009.

[41] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial," *JMIR Ment. Health*, vol. 4, no. 2, p. e19, Jun. 2017.

[42] L. M. Ritterband, F. Thorndike, L. A. Gonder-Frederick, J. C. Magee, E. T. Bailey, D. K. Saylor, and C. M. Morin, "Efficacy of an Internet-based behavioral intervention for adults with insomnia," *Arch. Gen. Psychiatry*, vol. 66, no. 7, pp. 692–698, Jul. 2009.

[43] D. F. Tolin, B. McGrath, L. R. Hale, D. N. Weiner, and R. Gueorguieva, "A multisite benchmarking trial of capnometry guided respiratory intervention for panic disorder in naturalistic treatment settings," *Appl. Psychophysiol. Biofeedback*, vol. 42, no. 1, pp. 51–58, Mar. 2017.

[44] NightWare, " NightWare improves sleep in patients with nightmare disorder," 2021. [Online]. Available: https://nightware.com/, Accessed on: Mar. 6, 2022.

[45] D. R. Christensen, R. D. Landes, L. Jackson, L. A. Marsch, M. J. Mancino, M. Chopra, and W. K. Bickel, "Adding an internet-delivered treatment to an efficacious treatment package for opioid dependence," *J. Consult. Clin. Psychol.*, vol. 82, no. 6, pp. 964–972, Dec. 2014.

[46] R. K. Merchant, R. Inamdar, and R. C. Quade, " Effectiveness of population health management using the propeller health asthma platform: A randomized clinical trial," *J. Allergy Clin. Immunol. Pract.*, vol. 4, no. 3, pp. 455–463, May/Jun. 2016.

[47] NuvoAir, "Respiratory care for COPD and asthma," 2021. [Online]. Available: https://www.nuvoair.com/, Accessed on: Mar. 6, 2022.

[48] Oleena, " Oleena registered digital therapeutic for cancer symptom management," 2020. [Online]. Available: https://oleena.com/ , Accessed on: Mar. 6, 2022.

[49] Insulia, "Insulia," 2017 [Online]. Available: https://insulia.com/, Accessed on: Mar. 6, 2022.

[50] M. F. Alwashmi, G. Mugford, W. Abu-Ashour, and M. Nuccio, "A digital diabetes prevention program (transform) for adults with prediabetes: Secondary analysis," *JMIR Diabetes*, vol. 4, no. 3, p. e13904, Jul. 2019.

[51] S. J. Athinarayanan, R. N. Adams, S. J. Hallberg, A. L. McKenzie, N. H. Bhanpuri, W. W. Campbell, J. S. Volek, S. D. Phinney, and J. McCarter, " Long-term effects of a novel continuous remote care intervention including nutritional ketosis for the management of type 2 diabetes: A 2-year non-randomized clinical trial," *Front. Endocrinol.*, vol. 10, p. 348, Jun. 2019.

[52] DarioHealth, " DarioHealth: Digital health solutions for chronic conditions," 2021. [Online]. Available: https://www.dariohealth.com/, Accessed on: Mar. 6, 2022.

[53] J. Downing, J. Bollyky, and J. Schneider, "Use of a connected glucose meter and certified diabetes educator coaching to decrease the likelihood of abnormal blood glucose excursions: The livongo for diabetes program," *J. Med. Internet Res.*, vol. 19, no. 7, p. e234, Jul. 2017.

[54] Constant Therapy Health, "Constant therapy," 2020. [Online]. Available: https://constanttherapyhealth.com/, Accessed on: Mar. 6, 2022.

[55] NovaVision, "NeuroEyeCoach," 2021. [Online]. Available: https://novavision.com/neuroeyecoach/, Accessed on: Mar. 6, 2022.

[56] BTS Bioengineering, "Nirvana-Virtual reality applied to neuromotor rehabilitation," 2021. [Online]. Available: https://www.btsbioengineering.com/nirvana/, Accessed on: Mar. 6, 2022.

[57] IMDRF SaMD Working Group, " Software as a medical device (SaMD): Key definitions," 2013. [Online]. Available: https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf, Accessed on: Mar. 6, 2022.

[58] U.S. Food & Drug Administration, " Premarket notification 510(k)," 2020. [Online]. Available: https://www.fda.gov/medical-devices/premarket-submissions/premarket-notification-510k, Accessed on: Mar. 6, 2022.

[59] U.S. Food & Drug Administration, " De novo classification request," 2019. [Online]. Available: https://www.fda.gov/medical-devices/premarket-submissions/de-novo-classification-request, Accessed on: Mar. 6, 2022.

[60] J. L. Johnston, S. S. Dhruva, J. S. Ross, and V. K. Rathi, "Early experience with the FDA's breakthrough devices program," *Nat. Biotechnol.*, vol. 38, no. 8, pp. 933–938, Jul. 2020.

[61] U.S. Food & Drug Administration, " Digital health software precertification (pre-cert) program," 2019. [Online]. Available: https://www.fda.gov/medical-devices/digital-health/digital-health-software-precertification-pre-cert-program, Accessed on: Mar. 6, 2022.

[62] U.S. Food & Drug Administration, " De novo classification request," 2016. [Online]. Available: https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/de-novo-classification-request, Accessed on: Mar. 6, 2022.

[63] M. Caffrey, "Digital health provider noom wins full CDC recognition for mobile, online applications," 2017. [Online]. Available: https://www.ajmc.com/newsroom/digital-health-provider-noom-wins-full-cdc-recognition-for-mobile-online-applications, Accessed on: Mar. 6, 2022.

[64] M. Caffrey, "Omada health receives full CDC recognition for diabetes prevention," 2018. [Online]. Available: https://www.ajmc.com/newsroom/omada-health-receives-full-cdc-recognition-for-diabetes-prevention, Accessed on: Mar. 6, 2022.

[65] A. Michaelides, C. Raby, M. Wood, K. Farr, and T. Toro-Ramos, "Weight loss efficacy of a novel mobile diabetes prevention program delivery platform with human coaching," *BMJ Open Diabetes Res. Care*, vol. 4, no. 1, p. e000264, Sep. 2016.

[66] D. H. Gustafson, F. M. McTavish, M. Y. Chih, A. K. Atwood, R. A. Johnson, M. G. Boyle, M. S. Levy, H. Driscoll, S. M. Chisholm, L. Dillenburg, A. Isham, and D. Shah, " A smartphone application to support recovery from alcoholism: A randomized clinical trial," *JAMA Psychiatry*, vol. 71, no. 5, pp. 566–572, May 2014.

[67] F. Naughton, S. Hopewell, N. Lathia, R. Schalbroeck, C. Brown, C. Mascolo, A. McEwen, and S. Sutton, "A context-sensing mobile phone app (Q sense) for smoking cessation: A mixed-methods study," *JMIR Mhealth Uhealth*, vol. 4, no. 3, p. e106, Sep. 2016.

[68] S. Goldstein, B. C. Evans, D. Flack, A. Juarascio, S. Manasse, F. Q. Zhang, and E. M. Forman, "Return of the JITAI: Applying a just-in-

time adaptive intervention framework to the development of m-health solutions for addictive behaviors," *Int. J. Behav. Med.*, vol. 24, no. 5, pp. 673–682, Oct. 2017.

[69] DarioHealth, "Diabetes management with blood glucose monitoring system," 2021. [Online]. Available: https://www.dariohealth.com/solutions/diabetes-management/, Accessed on: Mar. 6, 2022.

[70] D. Muoio, "Omada health expands digital platform with type 2, hypertension programs, new features," 2018. [Online]. Available: https://www.mobihealthnews.com/content/omada-health-expands-digital-platform-type-2-hypertension-programs-new-features, Accessed on: Mar. 6, 2022.

[71] Livongo Public Relations, "Livongo launches hypertension," 2018 [Online]. Available: https://www.mylivongo.com/Livongo-Launches-Hypertension/, Accessed on: Mar. 6, 2022.

[72] H. W. Rodbard, "Treating hypertension reduces CV risks in diabetes care," 2019. [Online]. Available: https://www.healio.com/news/cardiology/20190807/treating-hypertension-reduces-cv-risks-in-diabetes-care, Accessed on: Mar. 6, 2022.

[73] B. M. Y. Cheung and C. Li, "Diabetes and hypertension: Is there a common metabolic pathway?" *Curr. Atheroscler. Rep.*, vol. 14, no. 2, pp. 160–166, Apr. 2012.

[74] I. Nahum-Shani, S. N. Smith, B. J. Spring, L. M. Collins, K. Witkiewitz, A. Tewari, and S. A. Murphy, "Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support," *Ann. Behav. Med.*, vol. 52, no. 6, pp. 446–462, May 2018.

[75] J. G. Thomas and D. S. Bond, "Behavioral response to a just-in-time adaptive intervention (JITAI) to reduce sedentary behavior in obese adults: Implications for JITAI optimization," *Health Psychol.*, vol. 34S, pp. 1261–1267, Dec. 2015.

[76] M. S. Businelle, M a, D. E. Kendzor, S. G. Frank, D. J. Vidrine, and D. W. Wetter, "An ecological momentary intervention for smoking cessation: Evaluation of feasibility and effectiveness," *J. Med. Internet Res.*, vol. 18, no. 12, p. e321, Dec. 2016.

[77] E. T. Hébert, C. K. Ra, A. C. Alexander, A. Helt, R. Moisiuc, D. E. Kendzor, D. J. Vidrine, R. K. Funk-Lawler, and M. S. Businelle, "A mobile just-in-time adaptive intervention for smoking cessation: Pilot randomized controlled trial," *J. Med. Internet Res.*, vol. 22, no. 3, p. e16907, Mar. 2020.

[78] B. Spring, "Sense2stop: Mobile sensor data to knowledge," 2017. [Online]. Available: https://clinicaltrials.gov/ct2/show/study/NCT03184389, Accessed on: Mar. 6, 2022.

[79] L. Yardley, B. J. Spring, H. Riper, L. G. Morrison, D. H. Crane, K. Curtis, G. C. Merchant, F. Naughton, and A. Blandford, "Understanding and promoting effective engagement with digital behavior change interventions," *Am. J. Prev. Med.*, vol. 51, no. 5, pp. 833–842, Nov. 2016.

[80] N. Alshurafa, J. Jain, R. Alharbi, G. Iakovlev, B. Spring, and A. Pfammatter, "Is more always better?: Discovering incentivized mHealth intervention engagement related to health behavior trends" *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 4, p. 153, Dec. 2018.

[81] B. J. Fogg, "Creating persuasive technologies: An eight-step design process," in *Proc. 4th Int. Conf. Persuasive Technology*, Claremont, USA, 2009, pp. 44.

[82] L. M. Ritterband, F. Thorndike, D. J. Cox, B. Kovatchev, and L. A. Gonder-Frederick, "A behavior change model for internet interventions," *Ann. Behav. Med.*, vol. 38, no. 1, pp. 18–27, Aug. 2009.

[83] A. Schmidt, M. Beigl, and H. W. Gellersen, "There is more to context than location," *Comput. Graph.*, vol. 23, no. 6, pp. 893–901, Dec. 1999.

[84] B. Schilit, N. Adams, and R. Want, "Context-aware computing applications," in *Proc. 1st Workshop on Mobile Computing Systems and Applications*, Santa Cruz, USA, 1994, pp. 85–90.

[85] ArisGlobal, "Cloud based end-to-end drug development platform," 2020. [Online]. Available: https://www.arisglobal.com/, Accessed on: Mar. 6, 2022.

[86] Clario, "Endpoint technology services for clinical trial management," 2021. [Online]. Available: https://www.bioclinica.com/, Accessed on: Mar. 6, 2022.

[87] Advarra, "Advarra: Enabling safer, smarter, faster clinical research," 2018. [Online]. Available: https://www.advarra.com/, Accessed on: Mar. 6, 2022.

[88] DSG, "DSG Clinical trial software and data management solutions," 2022. [Online]. Available: https://dsg-us.com/index.html, Accessed on: Mar. 6, 2022.

[89] Mednet, "eClinical platform for clinical trials-mednet," 2019. [Online]. Available: https://www.mednetsolutions.com/, Accessed on: Mar. 6, 2022.

[90] Medidata, "Medidata-unified life science platform," 2021. [Online]. Available: https://www.medidata.com/, Accessed on: Mar. 6, 2022.

[91] Parexel Int. Corporation, "Clinical research organization (CRO) & biopharmaceutical services," 2020. [Online]. Available: https://www.parexel.com/, Accessed on: Mar. 6, 2022.

[92] SimpleTrials, "SimpleTrials-clinical trial management system," 2020. [Online]. Available: https://www.simpletrials.com/, Accessed on: Mar. 6, 2022.

[93] Veeva, "Veeva systems," 2019. [Online]. Available: https://www.veeva.com/, Accessed on: Mar. 6, 2022.

[94] Global Information, Inc., "Clinical trial management systems (CTMS)," 2019. [Online]. Available: https://m.giikorea.co.kr/report/go240200-clinicaltrial-management-systems-ctms.html, Accessed on: Mar. 6, 2022.

[95] Y. R. Park, Y. J. Yoon, H. Koo, S. Yoo, C. M. Choi, S. H. Beck, and T. W. Kim, "Utilization of a clinical trial management system for the whole clinical trial process as an integrated database: System development," *J. Med. Internet Res.*, vol. 20, no. 4, p. e103, Apr. 2018.

[96] IBM, "IBM Watson health-AI healthcare solutions," 2022. [Online]. Available: https://www.ibm.com/watson-health, Accessed on: Mar. 6, 2022.

[97] Oracle, "Siebel clinical trial management system-Oracle," 2020. [Online]. Available: https://www.oracle.com/industries/life-sciences/clinical-trial-management.html, Accessed on: Mar. 6, 2022.

[98] PRAPrism, "PRAprism," 2022. [Online]. Available: https://www.nextrials.com/, Accessed on: Mar. 6, 2022.

[99] Winchester Business Systems, "Winchester business systems-moving science forward," 2020. [Online]. Available: https://wbsnet.com/, Accessed on: Mar. 6, 2022.

[100] M. Rose, "Open data kit," 2013. [Online]. Available: https://opendatakit.org/, Accessed on: Mar. 6, 2022.

[101] D. Ferreira, V. Kostakos, and A. K. Dey, "AWARE: Mobile context instrumentation framework," *Front. ICT*, vol. 2, p. 6, Apr. 2015.

[102] S. M. Hossain, T. Hnat, N. Saleheen, N. J. Nasrin, J. Noor, B. J. Ho, T. Condie, M. Srivastava, and S. Kumar, "mCerebrum: A mobile sensing software platform for development and validation of digital biomarkers and interventions," in *Proc. 15th ACM Conf. Embedded Network Sensor Systems*, Delft, Netherlands, 2017, pp. 7.

[103] M. Rabbi, M. Philyaw-Kotov, J. Lee, A. Mansour, L. Dent, X. L. Wang, R. Cunningham, E. Bonar, I. Nahum-Shani, P. Klasnja, M. Walton, and S. Murphy, "SARA: A mobile app to engage users in health data collection," in *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing and Proc. ACM Int. Symp. Wearable Computers*, Hawaii, Maui, 2017, pp. 781–789.

[104] R. Eckhoff, N. Kizakevich, V. Bakalov, Y. Y. Zhang, S. Bryant, and M. A. Hobbs, "A platform to build mobile health apps: The personal health intervention toolkit (PHIT)," *JMIR MHealth Uhealth*, vol. 3, no. 2, p. e46, Jun. 2015.

[105] N. Kizakevich, R. Hubal, J. Brown, J. Lyden, J. Spira, R. Eckhoff, Y. Y. Zhang, S. Bryant, and G. Munoz, "PHIT for duty, a mobile approach for psychological health intervention," *Stud. Health Technol. Inform.*, vol. 181, pp. 268–272, Sep. 2012.

[106] N. Kizakevich, R. Eckhoff, J. Brown, S. J. Tueller, B. Weimer, S. Bell, A. Weeks, L. L. Hourani, J. L. Spira, and L. A. King, "PHIT for duty, a mobile application for stress reduction, sleep improvement, and alcohol moderation," *Mil. Med.*, vol. 183, no. suppl_1, pp. 353–363, Mar. 2018.

[107] J. Nakazawa, H. Tokuda, and T. Yonezawa, "Sensorizer: An architecture for regenerating cyber physical data streams from the web," in *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing and Proc. ACM Int. Symp. Wearable Computers*, Osaka,

Japan, 2015, pp. 1599–1606.

[108] J. B. Sun, S. Park, G. Jung, Y. Jeong, U. Lee, K. M. Chung, C. Lee, H. Kim, S. Ahn, A. Khandoker, and L. Hadjileontiadis, "BeActive: Encouraging physical activities with just-in-time health intervention and micro financial incentives," in *Proc. Symp. Emerging Research from Asia and on Asian Contexts and Cultures*, Honolulu, USA, 2020, pp. 17–20.

[109] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *IEEE Commun. Mag.*, vol. 48, no. 9, pp. 140–150, Sep. 2010.

[110] B. Cowley, M. Filetti, K. Lukander, J. Torniainen, A. Henelius, L. Ahonen, O. Barral, I. Kosunen, T. Valtonen, M. Huotilainen, N. Ravaja, and G. Jacucci, "The psychophysiology primer: A guide to methods and a broad review with a focus on human-computer interaction," *Found. Trends® Hum. Comput. Interact.*, vol. 9, no. 3–4, pp. 151–308, Nov. 2016.

[111] M. L. Lee and A. K. Dey, "Reflecting on pills and phone use: Supporting awareness of functional abilities for older adults," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, Vancouver, Canada, 2011, pp. 2095–2104.

[112] J. Lee, E. Walker, W. Burleson, M. Kay, M. Buman, and E. B. Hekler, "Self-experimentation for behavior change: Design and formative evaluation of two approaches," in *Proc. CHI Conf. Human Factors in Computing Systems*, Denver, USA, 2017, pp. 6837–6849.

[113] S. Taki, S. Lymer, C. G. Russell, K. Campbell, R. Laws, K. L. Ong, R. Elliott, and E. Denney-Wilson, "Assessing user engagement of an mHealth intervention: Development and implementation of the growing healthy app engagement index," *JMIR Mhealth Uhealth*, vol. 5, no. 6, p. e89, Jun. 2017.

[114] J. Brooke, "SUS: A 'quick and dirty' usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, I. L. McClelland, B. Weerdmeester, Eds. London, UK: CRC Press, 1996, pp. 189–194.

[115] A. M. Lund, "Measuring usability with the use questionnaire," *Usability Interface*, vol. 8, no. 2, pp. 3–6, Jan. 2001.

[116] H. L. O'Brien, Cairns, and M. Hall, "A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form," *Int. J. Hum. Comput. Stud.*, vol. 112, pp. 28–39, Apr. 2018.

[117] S. R. Stoyanov, L. Hides, D. J. Kavanagh, O. Zelenko, D. Tjondronegoro, and M. Mani, "Mobile app rating scale: A new tool for assessing the quality of health mobile apps," *JMIR Mhealth Uhealth*, vol. 3, no. 1, p. e27, Mar. 2015.

[118] M. Brhel, H. Meth, A. Maedche, and K. Werder, "Exploring principles of user-centered agile software development: A literature review," *Inf. Softw. Technol.*, vol. 61, pp. 163–181, May 2015.

[119] D. Ben-Zeev, E. A. Scherer, J. D. Gottlieb, A. J. Rotondi, M. F. Brunette, E. D. Achtyes, K. T. Mueser, S. Gingerich, C. J. Brenner, M. Begale, D. C Mohr, N. Schooler, Marcy, D. G. Robinson, and J. M. Kane, "mHealth for schizophrenia: Patient engagement with a mobile phone intervention following hospital discharge," *JMIR Ment. Health*, vol. 3, no. 3, p. e34, Jul. 2016.

[120] R. M. Martey, K. Kenski, J. Folkestad, L. Feldman, E. Gordis, A. Shaw, J. Stromer-Galley, B. Clegg, H. Zhang, N. Kaufman, A. N. Rabkin, S. Shaikh, and T. Strzalkowski, "Measuring game engagement: Multiple methods and construct complexity," *Simul. Gaming*, vol. 45, no. 4–5, pp. 528–547, Nov. 2014.

[121] E. Di Lascio, S. Gashi, and S. Santini, "Unobtrusive assessment of students' emotional engagement during lectures using electrodermal activity sensors," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 3, p. 103, Sep. 2018.

[122] A. Mathur, N. D. Lane, and F. Kawsar, "Engagement-aware computing: Modelling user engagement from mobile contexts," in *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing*, Germany, Heidelberg, 2016, pp. 622–633.

[123] M. Pielot, B. Cardoso, K. Katevas, J. Serrà, A. Matic, and N. Oliver, "Beyond interruptibility: Predicting opportune moments to engage mobile phone users," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 3, p. 91, Sep. 2017.

[124] M. Ghahramani, M. C. Zhou, and G. Wang, "Urban sensing based on mobile phone data: Approaches, applications, and challenges,"

[125] S. Imran, T. Mahmood, A. Morshed, and T. Sellis, "Big data analytics in healthcare-a systematic literature review and roadmap for practical implementation," *IEEE/CAA J. Automa. Sinica*, vol. 8, no. 1, pp. 1–22, Jan. 2021.

[126] J. Firth, J. Torous, J. Nicholas, R. Carney, A. Pratap, S. Rosenbaum, and J. Sarris, "The efficacy of smartphone-based mental health interventions for depressive symptoms: A meta-analysis of randomized controlled trials," *World Psychiatry*, vol. 16, no. 3, pp. 287–298, Oct. 2017.

[127] J. Firth, J. Torous, J. Nicholas, R. Carney, S. Rosenbaum, and J. Sarris, "Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials" *J. Affect. Disord.*, vol. 218, pp. 15–22, Aug. 2017.

[128] M. X. Cui, X. Y. Wu, J. F. Mao, X. Wang, and M. Nie, "T2DM self-management via smartphone applications: A systematic review and meta-analysis," *PLoS One*, vol. 11, no. 11, p. e0166718, Nov. 2016.

[129] Klasnja, E. B. Hekler, S. Shiffman, A. Boruvka, D. Almirall, A. Tewari, and S. A. Murphy, "Microrandomized trials: An experimental design for developing just-in-time adaptive interventions," *Health Psychol.*, vol. 34S, pp. 1220–1228, Dec. 2015.

[130] Li ao, Klasnja, A. Tewari, and S. A. Murphy, "Sample size calculations for micro-randomized trials in mHealth," *Statist. Med.*, vol. 35, no. 12, pp. 1944–1971, May 2016.

[131] W. Dempsey, Li ao, Klasnja, I. Nahum-Shani, and S. A. Murphy, "Randomised trials for the fitbit generation," *Significance*, vol. 12, no. 6, pp. 20–23, Dec. 2015.

[132] E. Littlewood, A. Duarte, C. Hewitt, *et al.*, "A randomised controlled trial of computerised cognitive behaviour therapy for the treatment of depression in primary care: The Randomised Evaluation of the Effectiveness and Acceptability of Computerised Therapy (REEACT) trial," *Health Technol. Assess.*, vol. 19, no. 101, Dec. 2015.

[133] M. J. Park and H. S. Kim, "Evaluation of mobile phone and internet intervention on waist circumference and blood pressure in post-menopausal women with abdominal obesity," *Int. J. Med. Inform.*, vol. 81, no. 6, pp. 388–394, Jun. 2012.

[134] S. Nundy, J. J. Dick, C. H. Chou, R. S. Nocon, M. H. Chin, and M. E. Peek, "Mobile phone diabetes project led to improved glycemic control and net savings for Chicago plan participants," *Health Aff.*, vol. 33, no. 2, pp. 265–272, Feb. 2014.

[135] F. Tsapeli and M. Musolesi, "Investigating causality in human behavior from smartphone sensor data: A quasi-experimental approach," *EPJ Data Sci.*, vol. 4, no. 1, p. 24, Dec. 2015.

[136] A. Mehrotra, F. Tsapeli, R. Hendley, and M. Musolesi, "MyTraces: Investigating correlation and causation between users' emotional states and mobile phone interaction," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 3, p. 83, Sep. 2017.

[137] J. Dallery, R. N. Cassidy, and B. R. Raiff, "Single-case experimental designs to evaluate novel technology-based health interventions," *J. Med. Internet Res.*, vol. 15, no. 2, p. e22, Feb. 2013.

[138] S. McDonald, F. Quinn, R. Vieira, N. O'Brien, M. White, D. W. Johnston, and F. F. Sniehotta, "The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: A systematic literature overview," *Health Psychol. Rev.*, vol. 11, no. 4, pp. 307–323, Dec. 2017.

[139] S. H. Jain, B. W. Powers, J. B. Hawkins, and J. S. Brownstein, "The digital phenotype," *Nat. Biotechnol.*, vol. 33, no. 5, pp. 462–463, May 2015.

[140] G. M. Harari, N. D. Lane, R. Wang, B. S. Crosier, A. T. Campbell, and S. D. Gosling, "Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges," *Perspect. Psychol. Sci.*, vol. 11, no. 6, pp. 838–854, Nov. 2016.

[141] T. R. Insel, "Digital phenotyping: Technology for a new science of behavior," *JAMA*, vol. 318, no. 13, pp. 1215–1216, 2017.

[142] A. K. Dey, "Understanding and using context," *Pers. Ubiquitous Comput.*, vol. 5, no. 1, pp. 4–7, Feb. 2001.

[143] A. Schmidt, M. Beigl, and H. W. Gellersen, "There is more to context than location," *Comput. Graph.*, vol. 23, no. 6, pp. 893–901, Dec. 1999.

IEEE/CAA J. Autom. Sinica, vol. 7, no. 3, pp. 627–637, May 2020.

[144] D. C. Mohr, M. Zhang, and S. M. Schueller, " Personal sensing: Understanding mental health using ubiquitous sensors and machine learning," *Ann. Rev. Clin. Psychol.*, vol. 13, pp. 23–47, May 2017.

[145] S. van Dantzig, G. Geleijnse, and A. T. van Halteren, " Toward a persuasive mobile application to reduce sedentary behavior," *Pers. Ubiquit. Comput.*, vol. 17, no. 6, pp. 1237–1246, Jul. 2013.

[146] S. Abdullah, M. Matthews, E. L. Murnane, G. Gay, and T. Choudhury, "Towards circadian computing: "early to bed and early to rise" makes some of us unhealthy and sleep deprived," in *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing*, Washington, Seattle, 2014, pp. 673–684.

[147] E. Thomaz, I. Essa, and G. D. Abowd, " A practical approach for recognizing eating moments with wrist-mounted inertial sensing," in *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing*, Osaka, Japan, 2015, pp. 1029–1040.

[148] R. Alam, J. Q. Gong, M. Hanson, A. Bankole, M. Anderson, T. Smith-Jackson, and J. Lach, " Motion biomarkers for early detection of dementia-related agitation," in *Proc. 1st Workshop on Digital Biomarkers*, Niagara Falls, USA, 2017, pp. 15–20.

[149] L. Canzian and M. Musolesi, "Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis," in *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing*, Osaka, Japan, 2015, pp. 1293–1304.

[150] Y. Huang, J. Q. Gong, M. Rucker, P. Chow, K. Fua, M. S. Gerber, B. Teachman, and L. E. Barnes, " Discovery of behavioral markers of social anxiety from smartphone sensor data," in *Proc. 1st Workshop on Digital Biomarkers*, Niagara Falls, USA, 2017, pp. 9–14.

[151] G. Mark, V. M. Gonzalez, and J. Harris, " No task left behind?: Examining the nature of fragmented work," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, Portland, USA, 2005, pp. 321–330.

[152] P. D. Adamczyk and B. P. Bailey, "If not now, when?: The effects of interruption at different moments within task execution," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, Vienna, Austria, 2004, pp. 271–278.

[153] J. Ho and S. S. Intille, "Using context-aware computing to reduce the perceived burden of interruptions from mobile devices," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, Bremen, Germany, 2005, pp. 909–918.

[154] W. Holland, " Statistics and causal inference," *J. Am. Stat. Assoc.*, vol. 81, no. 396, pp. 945–960, Dec. 1986.

[155] D. B. Rubin, " Causal inference using potential outcomes: Design, modeling, decisions," *J. Am. Stat. Assoc.*, vol. 100, no. 469, pp. 322–331, Feb. 2005.

[156] A. Boruvka, D. Almirall, K. Witkiewitz, and S. A. Murphy, "Assessing time-varying causal effect moderation in mobile health," *J. Am. Stat. Assoc.*, vol. 113, no. 523, pp. 1112–1121, Oct. 2018.

[157] T. C. Qian, Klasnja, and S. A. Murphy, "Linear mixed models with endogenous covariates: Modeling sequential treatment effects with application to a mobile health study," *Stat. Sci.*, vol. 35, no. 3, pp. 375–390, Aug. 2020.

[158] T. C. Qian, H. Yoo, Klasnja, D. Almirall, and S. A. Murphy, "Estimating time-varying causal excursion effects in mobile health with binary outcomes," *Biometrika*, vol. 108, no. 3, pp. 507–527, Sep. 2021.

[159] W. Van den Noortgate and Onghena, "The aggregation of single-case results using hierarchical linear models," *Behav. Anal. Today*, vol. 8, no. 2, pp. 196–209, Jan. 2007.

[160] E. J. Daza, " Causal analysis of self-tracked time series data using a counterfactual framework for n-of-1 trials," *Methods Inf. Med.*, vol. 57, no. 1, pp. e10–e21, Feb. 2018.

[161] M. A. Hernàn and J. M. Robins, *Causal Inference: What If.* Boca Raton: Chapman & Hall/CRC, 2020.

[162] R. Rosenbaum and D. B. Rubin, " The central role of the propensity score in observational studies for causal effects," *Biometrika*, vol. 70, no. 1, pp. 41–55, Apr. 1983.

[163] S. Sizemore and R. Alkurdi, "Matching methods for causal inference: A machine learning update," 2019. [Online]. Available: https://humboldt-wi.github.io/blog/research/applied_predictive_modeling_19/matching_methods/, Accessed on: Mar. 6, 2022.

[164] J. M. Robins and M. A. Hernan, "Estimation of the causal effects of time-varying exposures," in *Longitudinal Data Analysis*, G. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs, Eds. New York, USA: Chapman & Hall/CRC, 2008, pp. 553–597.

[165] J. Pearl, *Causality*. 2nd ed. Cambridge, USA: Cambridge University Press, 2009.

[166] C. W. J. Granger, " Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, Aug. 1969.

[167] G. Sugihara, R. May, H. Ye, C. H. Hsieh, E. Deyle, M. Fogarty, and S. Munch, " Detecting causality in complex ecosystems," *Science*, vol. 338, no. 6106, pp. 496–500, Sep. 2012.

[168] J. M. Twenge, G. N. Martin, and W. K. Campbell, " Decreases in psychological well-being among American adolescents after 2012 and links to screen time during the rise of smartphone technology," *Emotion*, vol. 18, no. 6, pp. 765–780, Sep. 2018.

[169] Z. Sarsenbayeva, G. Marini, N. van Berkel, C. Luo, W. W. Jiang, K. N. Yang, G. Wadley, T. Dingler, V. Kostakos, and J. Goncalves, "Does smartphone use drive our emotions or vice versa? A causal analysis," in *Proc. CHI Conf. Human Factors in Computing Systems*, Honolulu, USA, 2020, pp. 1–15.

[170] A. E. Yuan and W. Shou, " Data-driven causal analysis of observational time series in ecology," *bioRxiv*, 2021.

[171] T. Munzner, *Visualization Analysis and Design*. New York, USA: CRC Press, 2014.

[172] K. A. Cook and J. J. Thomas, "Illuminating the path: The research and development agenda for visual analytics," 2005. [Online]. Available: https://www.osti.gov/biblio/912515, Accessed on: Mar. 6, 2022.

[173] W. Aigner, S. Miksch, W. Muller, H. Schumann, and C. Tominski, "Visualizing time-oriented data—A systematic view," *Comput. Graph.*, vol. 31, no. 3, pp. 401–409, Jun. 2007.

[174] M. Brehmer, B. Lee, B. Bach, N. H. Riche, and T. Munzner, "Timelines revisited: A design space and considerations for expressive storytelling," *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 9, pp. 2151–2164, Sep. 2017.

[175] J. Walker, R. Borgo, and M. W. Jones, " TimeNotes: A study on effective chart visualization and interaction techniques for time-series data," *IEEE Trans. Vis. Comput. Graph.*, vol. 22, no. 1, pp. 549–558, Jan. 2016.

[176] R. Chen, F. Jankovic, N. Marinsek, L. Foschini, L. Kourtis, A. Signorini, M. Pugh, J. Shen, R. Yaari, V. Maljkovic, M. Sunga, H. H. Song, H. J. Jung, B. Tseng, and A. Trister, "Developing measures of cognitive impairment in the real world from consumer-grade multimodal sensor streams," in *Proc. 25th ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining*, Anchorage, USA, 2019, pp. 2145–2155.

[177] Y. Kim, E. Heo, H. Lee, S. Ji, J. Choi, J. W. Kim, J. Lee, and S. Yoo, "Prescribing 10, 000 steps like aspirin: Designing a novel interface for data-driven medical consultations," in *Proc. CHI Conf. Human Factors in Computing Systems*, Denver, USA, 2017, pp. 5787–5799.

[178] M. Sharmin, A. Raij, D. Epstien, I. Nahum-Shani, J. G. Beck, S. Vhaduri, K. Preston, and S. Kumar, " Visualization of time-series sensor data to inform the design of just-in-time adaptive stress interventions," in *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing*, Osaka, Japan, 2015, pp. 505–516.

[179] J. Polack Jr., S. T. Chen, M. Kahng, K. De Barbaro, R. Basole, M. Sharmin, and D. H. Chau, " Chronodes: Interactive multifocus exploration of event sequences," *ACM Trans. Interact. Intell. Syst.*, vol. 8, no. 1, p. 2, Mar. 2018.

[180] F. Birnbaum, D. Lewis, R. K. Rosen, and M. L. Ranney, "Patient engagement and the design of digital health," *Acad. Emerg. Med.*, vol. 22, no. 6, pp. 754–756, Jun. 2015.

[181] K. Skivington, L. Matthews, S. A. Simpson, Craig, J. Baird, J. M. Blazeby, K. A. Boyd, N. Craig, D. French, E. McIntosh, M. Petticrew, J. Rycroft-Malone, M. White, and L. Moore, " Framework for the development and evaluation of complex interventions: Gap analysis, workshop and consultation-informed update," *Health Technol. Assess.*, vol. 25, no. 57, pp. 1–132, Sep. 2021.

[182] S. Michie, M. Richardson, M. Johnston, C. Abraham, J. Francis, W. Hardeman, M. Eccles, J. Cane, and C. E. Wood, "The behavior change

technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions," *Ann. Behav. Med.*, vol. 46, no. 1, pp. 81–95, Aug. 2013.

[183] D. E. Conroy, C. H. Yang, and J. Maher, "Behavior change techniques in top-ranked mobile apps for physical activity," *Am. J. Prev. Med.*, vol. 46, no. 6, pp. 649–652, Jun. 2014.

[184] D. X. Chen, Z. J. Ding, C. G. Yan, and M. M. Wang, "A behavioral authentication method for mobile based on browsing behaviors," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 6, pp. 1528–1541, Nov. 2020.

[185] R. Orji, R. L. Mandryk, and J. Vassileva, "Towards a data-driven approach to intervention design: A predictive path model of healthy eating determinants," in *Proc. 7th Int. Conf. Persuasive Technology: Design for Health and Safety*, Linköping, Sweden, 2012, pp. 203–214.

[186] N. Alkış and T. T. Temizel, "The impact of individual differences on influence strategies," *Pers. Individ. Differ.*, vol. 87, pp. 147–152, Dec. 2015.

[187] M. Kaptein, Markopoulos, B. De Ruyter, and E. Aarts, "Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles," *Int. J. Hum. Comput. Stud.*, vol. 77, pp. 38–51, May 2015.

[188] W. B. Yue, Z. D. Wang, J. Y. Zhang, and X. H. Liu, "An overview of recommendation techniques and their applications in healthcare," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 4, pp. 701–717, Apr. 2021.

[189] S. E. Boslaugh, M. W. Kreuter, R. A. Nicholson, and K. Naleid, "Comparing demographic, health status and psychosocial strategies of audience segmentation to promote physical activity," *Health Educ. Res.*, vol. 20, no. 4, pp. 430–438, Aug. 2005.

[190] N. Askham, D. Cook, M. Doyle, H. Fereday, M. Gibson, U. Landbeck, R. Lee, C. Maynard, G. Palmer, and J. Schwarzenbach, "The six primary dimensions for data quality assessment," DAMA UK Working Group, UK, 2013.

[191] M. M. Breunig, H. P. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying density-based local outliers," in *Proc. ACM SIGMOD Int. Conf. Management of Data*, Dallas, USA, 2000, pp. 93–104.

[192] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, "Long short term memory networks for anomaly detection in time series," in *Proc. 23rd European Symp. Artificial Neural Networks, Computational Intelligence and Machine Learning*, Bruges, Belgium, 2015, pp. 89–94.

[193] A. Mehrotra, R. Hendley, and M. Musolesi, "NotifyMeHere: Intelligent notification delivery in multi-device environments," in *Proc. Conf. Human Information Interaction and Retrieval*, Glasgow, UK, 2019, pp. 103–111.

[194] D. Weber, A. Voit, P. Kratzer, and N. Henze, "In-situ investigation of notifications in multi-device environments," in *Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing*, Germany, Heidelberg, 2016, pp. 1259–1264.

[195] J. L. Hill, "Bayesian nonparametric modeling for causal inference," *J. Comput. Graph. Stat.*, vol. 20, no. 1, pp. 217–240, Mar. 2011.

[196] U. Shalit, F. D. Johansson, and D. A. Sontag, "Estimating individual treatment effect: Generalization bounds and algorithms," in *Proc. 34th Int. Conf. Machine Learning*, Sydney, Australia, 2017, pp. 3076–3085.

[197] A. M. Alaa, M. Weisz, and M. van der Schaar, "Deep counterfactual networks with propensity-dropout," in *Proc. 34th Int. Conf. Machine Learning*, Sydney, Australia, 2017, pp. 1–5.

[198] S. Wager and S. Athey, "Estimation and inference of heterogeneous treatment effects using random forests," *J. Am. Stat. Assoc.*, vol. 113, no. 523, pp. 1228–1242, Jun. 2018.

[199] J. Yoon, J. Jordan, and M. van der Schaar, "GANITE: Estimation of individualized treatment effects using generative adversarial nets," in *Proc. 6th Int. Conf. Learning Representations*, Vancouver, Canada, 2018, pp. 1–22.

[200] C. Shi, D. M. Blei, and V. Veitch, "Adapting neural networks for the estimation of treatment effects," in *Proc. 33rd Int. Conf. Neural Information Processing Systems*, Vancouver, Canada, 2019, pp. 225.

[201] B. Lim, A. Alaa, and M. van der Schaar, "Forecasting treatment responses over time using recurrent marginal structural networks," in *Proc. 32nd Int. Conf. Neural Information Processing Systems*, Montreal, Canada, 2018, pp. 7494–7504.

[202] I. Bica, A. M. Alaa, J. Jordon, and M. van der Schaar, "Estimating counterfactual treatment outcomes over time through adversarially balanced representations," in *Proc. 8th Int. Conf. Learning Representations*, Addis Ababa, Ethiopia, 2020, pp. 1–28.

[203] R. Q. Liu, C. C. Yin, and P. Zhang, "Estimating individual treatment effects with time-varying confounders," in *Proc. IEEE Int. Conf. Data Mining*, Sorrento, Italy, 2020, pp. 382–391.

[204] D. Brodbeck, R. Gasser, M. Degen, S. Reichlin, and J. Luthiger, "Enabling large-scale telemedical disease management through interactive visualization," *European Notes in Medical Informatics*, vol. 1, no. 1, pp. 1172–1177, Aug. 2005.

[205] A. Perer and D. Gotz, "Data-driven exploration of care plans for patients," in *Proc. CHI'13 Extended Abstracts on Human Factors in Computing Systems*, Paris, France, 2013, pp. 439–444.

[206] D. A. Rohani, M. Faurholt-Jepsen, L. V. Kessing, and J. E. Bardram, "Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: Systematic review," *JMIR Mhealth Uhealth*, vol. 6, no. 8, p. e165, Aug. 2018.

[207] D. Pach, A. A. Rogge, J. N. Wang, and C. M. Witt, "Five lessons learned from randomized controlled trials on mobile health interventions: Consensus procedure on practical recommendations for sustainable research," *JMIR Mhealth Uhealth*, vol. 9, no. 2, p. e20630, Feb. 2021.

[208] S. E. Lord, A. N. C. Campbell, M. F. Brunette, L. Cubillos, S. M. Bartels, W. C. Torrey, A. L. Olson, S. H. Chapman, J. A. Batsis, D. Polsky, E. V. Nunes, K. M. Seavey, and L. A. Marsch, "Workshop on implementation science and digital therapeutics for behavioral health," *JMIR Ment. Health*, vol. 8, no. 1, p. e17662, Jan. 2021.

[209] M. Gonzàlez, J. Lorés, and A. Granollers, "Enhancing usability testing through datamining techniques: A novel approach to detecting usability problem patterns for a context of use," *Inf. Softw. Technol.*, vol. 50, no. 6, pp. 547–568, May 2008.

[210] J. I. Gold, M. SooHoo, A. M. Laikin, A. S. Lane, and M. J. Klein, "Effect of an immersive virtual reality intervention on pain and anxiety associated with peripheral intravenous catheter placement in the pediatric setting: A randomized clinical trial," *JAMA Netw. Open*, vol. 4, no. 8, p. e2122569, Aug. 2021.

[211] L. T. Car, D. A. Dhinagaran, B. M. Kyaw, T. Kowatsch, S. Joty, Y. L. Theng, and R. Atun, "Conversational agents in health care: Scoping review and conceptual analysis," *J. Med. Internet Res.*, vol. 22, no. 8, p. e17158, Aug. 2020.

[212] Z. H. Shen, A. Elibol, and N. Y. Chong, "Understanding nonverbal communication cues of human personality traits in human-robot interaction," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 6, pp. 1465–1477, Nov. 2020.

[213] Z. T. Liu, M. Wu, W. H. Cao, L. F. Chen, J. Xu, R. Zhang, M. T. Zhou, and J. W. Mao, "A facial expression emotion recognition based human-robot interaction system," *IEEE/CAA J. Autom. Sinica*, vol. 4, no. 4, pp. 668–676, Sep. 2017.

[214] J. Rooksby, A. Morrison, and D. Murray-Rust, "Student perspectives on digital phenotyping: The acceptability of using smartphone data to assess mental health," in *Proc. CHI Conf. Human Factors in Computing Systems*, Glasgow, UK, 2019, pp. 425.

[215] H. Lee, S. Kang, and U. Lee, "Understanding privacy risks and perceived benefits in open dataset collection for mobile affective computing," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 6, no. 2, p. 61, Jul. 2022.

[216] J. Huh-Yoo, R. Kadri, and L. R. Buis, "Pervasive healthcare IRBs and ethics reviews in research: Going beyond the paperwork," *IEEE Pervas. Comput.*, vol. 20, no. 1, pp. 40–44, Jan./Mar. 2021.

[217] R. Chen, F. Jankovic, N. Marinsek, L. Foschini, L. Kourtis, A. Signorini, M. Pugh, J. Shen, R. Yaari, V. Maljkovic, M. Sunga, H. H. Song, H. J. Jung, B. Tseng, and A. Trister, "Developing measures of cognitive impairment in the real world from consumer-grade multimodal sensor streams," in *Proc. 25th ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining*, Anchorage, USA, 2019, pp. 2145–2155.

[218] A. Alorwu, S. Kheirinejad, N. van Berkel, M. Kinnula, D. Ferreira, A. Visuri, and S. Hosio, "Assessing MyData scenarios: Ethics, concerns, and the promise," in *Proc. CHI Conf. Human Factors in Computing*

Systems, Yokohama, Japan, 2021, pp. 209.

[219] L. F. Cranor, Guduru, and M. Arjula, "User interfaces for privacy agents," *ACM Trans. Comput. Hum. Interact.*, vol. 13, no. 2, pp. 135–178, Jun. 2006.

[220] H. Lee and U. Lee, "Dynamic consent for sensor-driven research," in *Proc. 13th Int. Conf. Mobile Computing and Ubiquitous Network*, Tokyo, Japan, 2021, pp. 1–6.

[221] F. Schaub, R. Balebako, A. L. Durity, and L. F. Cranor, "A design space for effective privacy notices," in *Proc. 11th USENIX Conf. Usable Privacy and Security*, Ottawa, Canada, 2015, pp. 1–17.

[222] Y. Y. Feng, Y. X. Yao, and N. Sadeh, "A design space for privacy choices: Towards meaningful privacy control in the internet of things," in *Proc. CHI Conf. Human Factors in Computing Systems*, Yokohama, Japan, 2021, pp. 64.

Uichin Lee is an Associate Professor in the School of Computing at KAIST. He received the B.S. in computer engineering from Chonbuk National University in 2001, the M.S. degree in computer science from KAIST in 2003, and the Ph.D. degree in computer science from UCLA in 2008. He continued his studies at UCLA as a postdoctoral research scientist (2008–2009) and then worked for Alcatel-Lucent Bell Labs as a member of technical staff till 2010. His major research areas are human-computer interaction (HCI), Ubiquitous computing (UbiComp), sensor data science (Internet of Things), and data visualization. He has been regularly serving as a program committee member of the key HCI conferences and journals such as ACM CHI, PACM IMWUT/ACM Ubicomp, and PACM HCI/ACM CSCW. He served as a program committee chair of ACM Ubicomp 2021, IEEE/IFIP WONS 2021, and general co-chairs of HCI Korea 2021 and IEEE ICMU 2021. He received the best paper awards at ACM CHI'16, AAAI ICWSM'13, IEEE CCGrid'11, and IEEE PerCom'07 and an impact award from IEEE IoT Forum'19.

Gyuwon Jung received the B.S. degree in chemistry from KAIST in 2016, and an M.S. degree in knowledge service engineering from KAIST in 2018. He is currently pursuing the Ph.D. degree in the School of Computing at KAIST. His research interests are behavioral change support system, digital health, and human-data interaction.

Eun-Yeol Ma received the B.S. degree in 2017 and M.S. degree in 2019 in industrial engineering from KAIST. He is currently pursuing the Ph.D. degree in the Department of Industrial and Systems Engineering, KAIST. His research interests include machine learning, applied statistics, and causal inference.

Jin San Kim is a graduate student in the Department of Electrical & Computer Engineering at Inha University. He received the B.S. degree in computer engineering from Inha university in 2020. His research interests include big data platform, data science, and mobile computing.

Heepyung Kim received the B.S. degree in biomedical engineering form Jeonbuk National University, in 2007, and the master and Ph.D. degrees in biomedical engineering form Hanyang University, in 2010 and 2016. he was a Postdoctoral Fellow with KAIST Institute for Health Science and Technology of KAIST, From 2016 to 2019. he is currently Research Professor of KAIST Institute for Health Science and Technology of KAIST. His current research interests include bio & speech signal processing, biomedical engineering, digital phenotyping, digital healthcare, and mental health.

Jumabek Alikhanov received the B.S. degree from Tashkent University of Information Technology in 2014, and the M.E. degree from Inha University, in 2017. He is currently pursuing the Ph.D. degree with the Department of Computer Science and Information Engineering, Inha University. His research interests include machine learning and its applications on computer vision, natural language processing, sensor data science and information security.

Youngtae Noh is an Associate Professor in Energy AI, KENTECH. He received the B.S. degree in computer science from Chosun University in 2005, an M.S. degree in information and communication from Gwangju Institute of Science Technology (GIST) in 2007, and the Ph.D. degree in computer science at the University of California, Los Angeles (UCLA) in 2012. Before joining KENTECH, he worked as an Associate Professor at Inha University (2015–2022). Moreover, he has also worked at Cisco Systems as a Staff Member (2012–2014). His research areas include mobile/pervasive computing, mobile systems, mobile data science, mobile-HCI, cloud computing, data center networking, wireless networking, and future Internet.

Heeyoung Kim received the B.S. and M.S. degrees in industrial engineering from KAIST, an M.S. degree in statistics and the Ph.D. degree in industrial engineering from the Georgia Institute of Technology. She is an Associate Professor with the Department of Industrial and Systems Engineering, KAIST. She was a Senior Member of Technical Staff with AT&T Laboratories. Her research interests include applied statistics and machine learning.