frontiers | Frontiers in Psychiatry

Check for updates

The Feasibility and Utility of Harnessing Digital Health to Understand Clinical Trajectories in Medication Treatment for Opioid Use Disorder: D-TECT Study Design and Methodological Considerations

Lisa A. Marsch[1]*, Ching-Hua Chen[2], Sara R. Adams[3], Asma Asyyed[4], Monique B. Does[3], Saeed Hassanpour[1,5], Emily Hichborn[1], Melanie Jackson-Morris[3], Nicholas C. Jacobson[1,5], Heather K. Jones[3], David Kotz[1,6], Chantal A. Lambert-Harris[1], Zhiguo Li[2], Bethany McLeman[1], Varun Mishra[7,8], Catherine Stanger[1], Geetha Subramaniam[9], Weiyi Wu[1,5] and Cynthia I. Campbell[3,10]

[1] Center for Technology and Behavioral Health, Geisel School of Medicine, Dartmouth College, Lebanon, NH, United States, [2] Center for Computational Health, International Business Machines (IBM) Research, Yorktown Heights, NY, United States, [3] Division of Research Kaiser Permanente Northern California, Oakland, CA, United States, [4] The Permanente Medical Group, Northern California, Addiction Medicine and Recovery Services, Oakland, CA, United States, [5] Department of Biomedical Data Science, Geisel School of Medicine, Dartmouth College, Lebanon, NH, United States, [6] Department of Computer Science, Dartmouth College, Hanover, NH, United States, [7] Khoury College of Computer Sciences, Northeastern University, Boston, MA, United States, [8] Department of Health Sciences, Bouvé College of Health Sciences, Northeastern University, Boston, MA, United States, [9] Center for Clinical Trials Network, National Institute on Drug Abuse, Bethesda, MD, United States, [10] Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, San Francisco, CA, United States

Introduction: Across the U.S., the prevalence of opioid use disorder (OUD) and the rates of opioid overdoses have risen precipitously in recent years. Several effective medications for OUD (MOUD) exist and have been shown to be life-saving. A large volume of research has identified a confluence of factors that predict attrition and continued substance use during substance use disorder treatment. However, much of this literature has examined a small set of potential moderators or mediators of outcomes in MOUD treatment and may lead to over-simplified accounts of treatment non-adherence. Digital health methodologies offer great promise for capturing intensive, longitudinal ecologically-valid data from individuals in MOUD treatment to extend our understanding of factors that impact treatment engagement and outcomes.

Methods: This paper describes the protocol (including the study design and methodological considerations) from a novel study supported by the National Drug Abuse Treatment Clinical Trials Network at the National Institute on Drug Abuse (NIDA). This study (D-TECT) primarily seeks to evaluate the feasibility of collecting ecological momentary assessment (EMA), smartphone and smartwatch sensor data, and social media data among patients in outpatient MOUD treatment. It secondarily seeks to examine the utility of EMA, digital sensing, and social media data (separately and

compared to one another) in predicting MOUD treatment retention, opioid use events, and medication adherence [as captured in electronic health records (EHR) and EMA data]. To our knowledge, this is the first project to include all three sources of digitally derived data (EMA, digital sensing, and social media) in understanding the clinical trajectories of patients in MOUD treatment. These multiple data streams will allow us to understand the relative and combined utility of collecting digital data from these diverse data sources. The inclusion of EHR data allows us to focus on the utility of digital health data in predicting objectively measured clinical outcomes.

Discussion: Results may be useful in elucidating novel relations between digital data sources and OUD treatment outcomes. It may also inform approaches to enhancing outcomes measurement in clinical trials by allowing for the assessment of dynamic interactions between individuals' daily lives and their MOUD treatment response.

Clinical Trial Registration: Identifier: NCT04535583.

INTRODUCTION

Across the U.S., the prevalence of opioid use disorder (OUD) and the rates of opioid overdoses have risen precipitously in recent years. Drug overdose has been called a "modern plague" (1) and is the leading cause of death of Americans under age 50, having surpassed peak death rates from gun violence, HIV, and car crashes (1, 2). Over 100,000 Americans died from a drug overdose from May 2020 to April 2021 (3). This dramatic spike in OUD has also been accompanied by marked increases in injection-related infections (including infective endocarditis and Hepatitis C) (4–7), babies born with Neonatal Opioid Withdrawal Syndrome (8) and healthcare and criminal justice costs (9).

Several effective medications for OUD (MOUD) have been shown to be life-saving including buprenorphine, methadone, and naltrexone products (10, 11), and to greatly increase opioid abstinence, reduce HIV/infectious disease risk behavior, and reduce criminality. Greater MOUD retention is associated with the most positive treatment outcomes (12–16). However, over 50% of patients receiving MOUD dropout of treatment within 3–6 months after treatment initiation (17–21), falling short of the longer threshold of treatment shown to offer sustained benefit (22, 23). Additionally, given the chronic relapsing nature of the disease of addiction, and inconsistent compliance with MOUD, many individuals continue to engage in opioid use during treatment, increasing the risk of overdose (24, 25).

Many factors predict attrition and continued substance use during substance use disorder (SUD) treatment (26, 27), including, stress, mental health comorbidities, continued exposure to high-risk social networks or contexts, and the neurobiology underpinning addiction. However, this literature has examined a limited set of predictors of outcomes in MOUD treatment and may not reflect a comprehensive understanding of treatment non-adherence (28). Further, treatment engagement is typically evaluated via structured clinical assessments conducted on an episodic basis and may not reflect factors in individuals' daily lives that impact their OUD treatment trajectories. Thus, there is tremendous opportunity to more frequently and extensively examine factors that impact individuals' clinical trajectories in MOUD in real time.

Digital methodologies offer great promise for capturing intensive, longitudinal ecologically-valid data from individuals receiving MOUD to extend our understanding of factors that impact treatment engagement and outcomes (29, 30). In particular, the use of digital devices such as smartphones or wearables that measure individuals' health-related behavior (sometimes referred to as "digital phenotyping") (31) has the potential to provide personalized health care resources. The ubiquity of digital devices and the explosion of "big data" analytics enable the collection and interpretation of enormous amounts of rich data about everyday behavior. This includes the use of digital devices to implement "ecological momentary assessment" (EMA) (32) in which individuals are asked to respond to brief queries on their mobile devices (assessing, for example, craving, mood, withdrawal symptoms, and pain). It also includes passive sensing data collected via sensors embedded in smartphones and/or wearable sensing devices such as smartwatches that provide information about the wearer's health (e.g., heart rate and heart-rate variability measured via wearable photoplethysmography), behavior (e.g., social contact via calls, texts and app use), and environment (e.g., location type via GPS) (33). And, it includes social media data that individuals produce (e.g., the images and the texts they post).

A rapidly growing literature is underscoring the utility of such digital health data-driven approaches to understanding human behavior (34–37). Digitally-derived data may similarly reveal new insights into the temporal dynamics between moderators and mediators of MOUD treatment outcomes. Such data may complement and extend data captured via structured clinical assessments and provide a more comprehensive understanding of each individual's course of treatment. And these data, in turn,

may increase our ability to develop more potent and personalized treatment models for OUD.

The developing literature on the application of digital health to understanding individuals' trajectories in SUD treatment has shown promise. One study that used EMA to identify predictors of substance use among adults after an initial episode of SUD treatment showed high EMA completion rates (81%) and identified specific substance use patterns, negative affect and craving as predictors of substance use (38). EMA research has also demonstrated differing relationships between drug triggers (e.g., exposure to drug cues or mood changes) and different types of drug use. Specifically, drug triggers increased for hours before cocaine use events but not before heroin use events (39). And, among smokers trying to quit, smoking lapses were associated with increases in negative mood for many days (and not just hours) (40).

Additionally, EMA research with adults in MOUD treatment demonstrated a stronger relationship between craving and drug use events than between stress and drug use events (41). EMA-assessed momentary pain has been shown to be indirectly associated with illicit opioid use via momentary opioid craving (42). Further, MOUD treatment dropout has been shown to be more likely among individuals who report more "hassles", higher levels of cocaine craving, lower levels of positive mood, a recent history of emotional abuse, and a recent history of being bothered frequently by psychological problems. It is noteworthy that none of those factors predicted individuals' non-compliance with completing EMA (43). Other EMA research revealed that patients in MOUD treatment who share similar patterns of drug use (frequent opioid use, frequent cocaine use, frequent dual use of opioids and cocaine, sporadic drug use, or infrequent drug use) tended to have similar psychological processes preceding drug use events (44).

Less research has focused on the utility of passively collected sensing data or social media data in predicting substance use. One study used GPS data from phones to assess exposure to visible signs of environmental disorder and poverty among adults in outpatient MOUD treatment. That study provided a proof of concept that digitally-captured environmental data could predict drug craving and stress 90 min into the future (45). Another study with adults in outpatient MOUD, focused on passive assessment of stress, showed that the duration of a prior stress episode predicts the duration of the next stress episode and that stress in the mornings and evenings is lower than during the day (46). And another study demonstrated that deep-learning analytic approaches applied to social media data may be useful in identifying potential substance use risk behavior, such as alcohol use (47).

Overall, these findings have provided some new insights into how data collected in naturalistic settings may enhance an understanding of risk profiles among individuals in SUD treatment. Nonetheless, the breadth of factors evaluated to date has been limited, and most digital health studies conducted with populations in SUD treatment have relied exclusively on self-reported clinical outcomes (with limited focus on objective metrics such as urine screens, medication fills, and clinical visits).

Additionally, most studies have solely sought to predict substance use events.

This paper describes the protocol (including the study design and methodological considerations) from a novel study supported by the National Drug Abuse Treatment Clinical Trials Network (CTN) at the National Institute on Drug Abuse (NIDA). This study, referred to as "Harnessing Digital Health to Understand Clinical Trajectories of Opioid Use Disorder" (D-TECT; CTN-0084-A2) primarily seeks to evaluate the feasibility of collecting EMA, digital sensing and social media data among patients in outpatient MOUD treatment. It secondarily seeks to examine the utility of EMA, digital sensing, and social media data (separately and compared to one another) in predicting MOUD treatment retention, opioid use events, and medication adherence [as captured in Electronic Health Records (EHR), medical claims, and EMA data]. This is the first project to include all three sources of digitally derived data (EMA, sensing and social media) in understanding the clinical trajectories of patients in MOUD treatment. Multiple data streams will allow us to understand the relative and combined utility of collecting digital data from these diverse data sources. The inclusion of EHR data allows us to focus on the utility of digital health data in predicting objectively measured clinical outcomes.

METHODS AND ANALYSIS

Overview of Study Design

Individuals with OUD will be recruited for the study from among patients who are active in outpatient MOUD treatment with buprenorphine medication for at least 2 weeks at one of four Addiction Medicine Recovery Services (AMRS) programs at Kaiser Permanente Northern California (KPNC). Once it is confirmed that eligibility criteria are met, each participant will provide electronic informed consent and complete the baseline assessment by phone. The baseline appointments will take ∼2.0 h to complete, which will be done in two to three visits. Participants will be asked to wear a smartwatch and carry a smartphone (a study-supplied one or their own) that will passively collect sensor data. They will be asked to actively respond to EMA prompts through a smartphone 3 times daily and to self-initiate EMA responses daily if substance use occurred over the 12-week study. For those who consent to the optional social media component, social media data will be downloaded by the participant directly from the social media platform to a secure server using a remote desktop at the beginning of the study and again at the end of the study. EHR data extraction will occur at ∼16 weeks after the full study is completed and will collect data 12 months prior to EMA start (the date the participant began receiving EMA prompts) through 12-weeks after EMA start (84 days after the EMA start date). A follow-up assessment (∼45 min in length) will occur by phone ∼12 weeks after EMA start. A graphic overview of the study phases is presented in **Figure 1**.

Study Sites and Rationale for Site Selection

KPNC is a large, integrated health care delivery system with 4.3 million members, providing care through commercial

FIGURE 1 | Study phases.
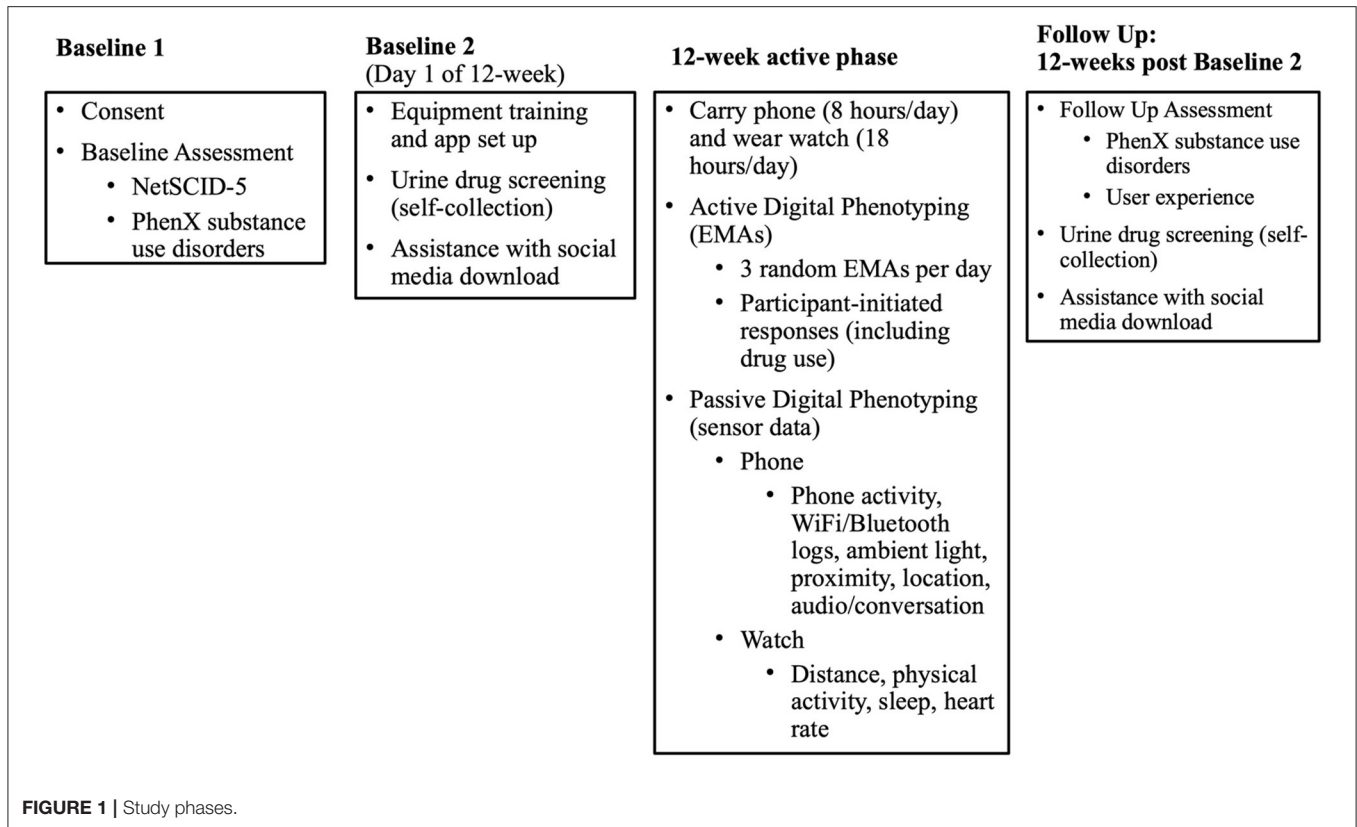
plans, Medicare, Medicaid, and health insurance exchanges. It is comprised of a racially and socioeconomically diverse membership and is generally representative of the region's population with access to care. KPNC was selected based on its ability to (1) provide access to individuals who are prescribed buprenorphine for OUD and (2) provide access to EHR data on treatment retention, medication adherence, and service utilization. KPNC maintains a data repository, the Virtual Data Warehouse, which has combined EHR data (e.g., demographics, membership, diagnoses, service utilization, pharmacy, lab data) with several other data sources, including medical claims data (e.g., non-Kaiser pharmacy data).

The AMRS programs at KPNC offer a broad range of services, including prescribing buprenorphine for OUD, medical services, group and individual therapy, and family therapy. Staffing includes physicians, therapists, medical assistants, nurses, and social workers.

Participants

Participants will include individuals aged 18 years or older across all racial and ethnic categories. Eligibility criteria include: active in KPNC outpatient treatment and prescribed buprenorphine for OUD for the past 2 weeks (and attended at least one visit at AMRS in past 35 days); $\geq$18 years old; capable of understanding and speaking English; able to participate in the full duration of the study (12 weeks); have an active email account and willing to provide its address to researchers; permit access to EHR data; willing to carry and use a personal or study-provided smartphone

for 12 weeks; and willing to wear a smartwatch continuously (except during pre-defined activities such as showering) for 12 weeks. Individuals will be excluded if they are: unwilling or unable to provide informed consent; currently in jail, prison or other overnight facility as required by court of law or have pending legal action that could prevent participation in study activities. We expect to recruit 50–75 participants.

Study Procedures
Recruitment and Screening

Potentially eligible patients will be initially identified from EHR data as meeting study criteria; eligibility is further confirmed through chart review. Eligible patients will be sent an invitational recruitment letter through a secure email message. Within approximately a week of mailing recruitment letters (or sending a secure email message), research staff will contact participants by phone to determine if they are interested and eligible, using the IRB-approved recruitment script, verbal consent form and final screening questions. If the individual is interested and eligible, research staff will schedule them for an initial baseline phone appointment and email them consent documents for their review before their appointment.

Informed Consent

Informed consent will be obtained by phone and documented online using an electronic signature. Each participant will be asked to pass a brief consent quiz to document comprehension of the study activities. The research staff will obtain authorization

from participants for use of protected health information, such as their EHR and medical claims data.

Baseline

The baseline process will be conducted in two or three phone appointments: the first appointment will consist of informed consent and the baseline assessment (Baseline 1), and the second (and third, if necessary) appointment (Baseline 2/3) will consist of a urine drug screen, setting up study devices, installing study applications ("apps"), and learning to use devices and apps (**Figure 1**).

The baseline assessment consists of interviewer-administered measures (described below) examining participant characteristics, current substance use (e.g., tobacco, alcohol, opioids, and other drugs), substance use and mental health disorders, and the impact of the COVID-19 pandemic. Once the first baseline appointment is completed, participants will be mailed a urine drug screen kit, a smartwatch and study smartphone (if applicable) and technology training documentation.

Once the equipment is received, research staff will schedule a second phone appointment with each participant to review the urine collection and technology training documentation. Research staff will walk through the set-up, use and care of the smartphone and smartwatch, installation of the study app and the Garmin Connect app (described below) if they are not already installed on a study-provided phone, as well as instructions for initiating and completing the daily EMA surveys. Research staff will also instruct participants to collect a urine sample and upload results.

Participants who consented to the social media part of the study will also receive instruction on how to request and download their social media data.

Active Study Phase (12-Weeks)

During the active 12-week study phase, the research staff will monitor participant compliance using a custom dashboard (e.g., their EMA completion rate and whether they carry the study phone and wear the study watch). In the first 2 weeks, participants will be followed closely. If after a 48-h period there are no EMA data, and/or no phone carry time data, and/or no watch wear data, research staff will follow-up directly with the participant via phone, text, and/or email to encourage the participant to continue their participation and/or troubleshoot any problems that may arise with the smartphone, smartwatch, and/or study app. All participants will have a 1-week check-in appointment via phone with research staff to review the participant's experience, review data collection over the past week, and answer any questions or resolve any technical issues with the study devices (regardless of device carry/wear time compliance or EMA completion rate). Thereafter, research staff will send weekly check-in texts or phone calls unless there is a 48-h period of no EMA data, and/or no phone carry time data, and/or no watch data. In those instances, the research staff will attempt to reach the participant by phone, text, and/or email. The research staff will make up to three contact attempts prior to engaging alternate contact(s).

Engagement in Research

To be considered engaged in the study, an individual must respond to a minimum number of EMA prompts (complete at least 2/3 of the EMA surveys per day on 7 out of the first 14 days) and record at least 8 h of smartphone/smartwatch sensor data per day on 7 out of the first 14 days of study participation. If an individual does not meet the engagement criterion and is non-responsive to research staff outreach in the first 14 days of study participation, then the individual will be considered a "non-engager" and the study team will continue to recruit until the targeted sample size is met. Non-engagers will not be withdrawn from the study, as we will attempt to collect all possible data from all participants.

Follow-Up

A follow-up assessment will be completed by phone ~12 weeks post-EMA start. Research staff will administer an interviewer-based assessment to measure current substance use, participant experience with the study devices, treatment utilization, reasons for drop out (if appropriate), employment, insurance coverage, medication use/dose (if applicable), and overdose (if applicable). Participants will be mailed a urine drug test kit and asked to collect another urine sample and upload test results. Participants who consented to the social media part of the study will be asked to request and download their social media data a second time.

Description of Measures

A summary of the clinical assessments and digital health assessments to be conducted in this study are reflected in **Figures 2**, **3**, respectively. Brief descriptions of each of these measures is provided below.

Prisoner Status Assessment: An individual's prisoner status must be assessed for each participant at each separate encounter, as this study will not apply for Office of Human Research Protection (OHRP) Prisoner Certification. An **Inclusion/Exclusion** form will be used to obtain information on inclusion and exclusion criteria to document eligibility. **Locator Form.** A locator form is used to obtain information at baseline and each contact to assist in finding participants throughout the study. This form collects the participant's current address, email address, and phone numbers. **PhenX Substance Abuse and Addiction Core Tier 1 (PhenX Core Tier 1).** The PhenX Core Tier 1 is a part of the Substance Abuse and Addiction Collections (48) that are being adopted across multiple studies funded by NIDA. This study will use the following subset of measures from the Core Tier 1: demographics (age, ethnicity, gender, race, current educational attainment, current employment status, and current marital status) and current substance use (tobacco, alcohol, and drugs) (49). The demographics (except for employment status) will only be collected at baseline, while current substance use and current employment status will be collected at baseline and follow-up. **The PhenX Substance Abuse and Addiction Core Tier 2** (PhenX Core Tier 2). The PhenX Core Tier 2 is a complementary set of 8 measures to the PhenX Core Tier 1 (i.e., annual family income, child-reported parental education attainment, family history of substance use problems,

Study Visit	Measures	Data Collection Method
Screening	• Verbal consent • Prisoner Status Assessment (repeated at each study visit) • Inclusion/Exclusion Form	• Interview
Baseline 1-3	• Prisoner Status Assessment • Electronic Consent (eConsent) • Electronic Health Insurance Portability and Accountability Act (eHIPAA) Authorization • Consent Quiz • Locator Form (completed at each call) • Phenotypes and eXposures project measures (PhenX) Core Tier 1 (demographics & current substance use) • PhenX Core Tier 2 (annual family income) • National Survey on Drug Use and Health (NSDUH) (annual family income)* • Baseline Assessment - COVID-19 (impacts due to the pandemic) • NetSCID-5 (computer-based Structured Clinical Interview for DSM-5 – research diagnosis) • Urine Drug Screen (UDS) (recent use)	• Interview
	• Social Media (images/texts of postings & comments, reactions, date/time)	• Download (from social media platform via remote desktop)
Follow-up	• Prisoner Status Assessment • PhenX Core Tier 1 (current substance use) • Follow-up Assessment (participant experience; employment; treatment utilization; insurance coverage) • UDS (recent use)	• Interview
	• Social Media (images/texts of postings & comments, reactions, date/time)	• Download (from social media platform via remote desktop)
	• Electronic Health Record (EHR)	• Extraction**
	• Medical Claims	• Extraction**

FIGURE 2 | Table of study assessments. *NSDUH is only collected for participants who are unsure of the total family income. We will use as subset of questions to determine which income category best characterizes total combined family income. **EHR/Medical Claims data will be extracted by the data analyst ~16 weeks after completion of study and includes data 12 months prior to EMA start through 12 weeks post-EMA start.

household roster-relationships, internalizing, externalizing, and substance use disorders screener, occupation/occupational history, peer/partner substance use and tolerance of substance use, and social networks) (50). This study will only use a subset of questions from the Annual Family Income measure to get an estimate of total income of all family members (49). If the participant is unsure of the total family income, then we will use a subset of questions from the Substance Abuse and Mental Health Services Administration's National Survey on Drug Use and Health (NSDUH) survey to determine which income

category best characterizes total combined family income (51). These measures will be securely, electronically stored in a REDCap database.

NetSCID-5: The Structured Clinical Interview for DSM-5 (SCID-5) is a semi-structured interview designed to assess substance use and mental health diagnoses (52). This study will use an electronic version of the SCID-5, the NetSCID-5, developed by TeleSage. The TeleSage NetSCID-5 is fully licensed by the American Psychiatric Association and has been validated (53). TeleSage has the capability of customizing

Device	Measures	Data Collection Method
Smartwatch	• Ambulatory Physiological Assessment using Mobile sensing (distance; physical activity; sleep; heart rate)	• Wearable*
Smartphone	• Passive Assessment using Mobile Sensors (phone activity; WiFi/Bluetooth logs; ambient light; proximity; location; audio/conversation)	• Carryable**
	• Ecological Momentary Assessment (EMA) (including self-report of drug use)	• Self-report***
	• Momentary Self-Regulation Scale 12-item	• Self-report***

FIGURE 3 | Active study phase digital health assessments. *Wear smartwatch at least 18 hours per day, data transmitted real-time for 12 weeks. **Carry smartphone at least 8 hours per day, data transmitted real-time for 12 weeks. ***3 times per day for 12 weeks.

the NetSCID-5 measure, and modules relevant for this study include: bipolar I disorder, major depressive disorder, panic disorder, social anxiety disorder, generalized anxiety disorder, posttraumatic stress disorder, adult attention deficit hyperactivity disorder, alcohol use disorder, and other use disorders [cannabis, stimulants/cocaine, opioids, phenylcyclohexyl piperidine (PCP), other hallucinogens, inhalants, sedative-hypnotic-anxiolytic, and other/unknown]. The TeleSage NetSCID-5 will be administered by research staff who are trained and credentialed to conduct this diagnostic assessment.

Urine Drug Screen: Urine drug screen kits will be mailed to participants, and participants will be asked to collect a urine sample and then record and upload its results using a secure system (e.g., REDCap) at first baseline appointment and at follow-up. All urine specimens are collected using CLIA-Waived and FDA-approved one-step multi-drug screen test cups following the manufacturer's recommended procedures. The study will use the DrugConfirm Advance Urine Drug Test Kit that screens for: alcohol, amphetamine, barbiturate, buprenorphine, benzodiazepine, cocaine, fentanyl, MDMA (ecstasy/molly), methamphetamine, methadone, morphine 300 ng/mL, oxycodone, phencyclidine (PCP), tramadol and delta-9-tetrahydrocannabinol (THC).

In order to reduce risks of substituted or adulterated urine samples, the research staff will conduct the study urine drug screens in real-time (i.e., reviewing collection instructions and walking the participant through the entire process via phone) during the Baseline 2 phone appointment. Participants will send a photo of the temperature strip to research staff immediately after they produce the sample to ensure the temperature is within the specified valid temperature range of 90°-100°F. Additionally, we will ask participants take photos of the test result strips and send the photos of the results securely to research staff in

real-time (while the participant is still on the phone). Research staff will review the photos that are sent to ensure that the results captured within the photos are legible and not blurry or otherwise indecipherable.

Digital Health Technology
Ambulatory Physiological Assessment Using Mobile Sensors

We will develop a smartphone application ("study app") for both Android and iOS devices. The study app can sense and store contextual information about a participant, e.g., location, physical activity (step count), conversation duration and count (non-identified audio information such as segments of silence, and speech features such as pitch control and voice quality), app usage, call/text, screen on/off, phone lock/unlock, and phone notifications (54). Features will be derived from the raw sensor streams to create multiple relevant contextual variables. This custom application will be installed directly on the study-provided smartphone (Moto G7 Power and/or Moto G Power) or on a participant's smartphone if they have a compatible phone (iPhones, running iOS 12 or higher or Android devices running Android 8.0 or higher with at least 2.5 GB RAM and 4 GB of available storage).

In addition to the smartphone, participants will be provided with a smartwatch (Garmin Vivosmart 4). The participants will be asked to wear the smartwatch continuously (except during pre-determined exception periods, such as when the participants are showering or charging the device). The Garmin Vivosmart 4 smartwatch is comfortable, lightweight, and has a long battery life of up to 7 days (55) and an easy-to-use interface. The device can continuously collect and track a variety of sensor data in the background, as long as the user is wearing the device. The data from the wearable is synced directly with Garmin cloud servers,

using the "Garmin Connect" application installed on the phone, and we will not have direct access to the raw sensor data. We will use the Garmin Health Connect API to get various health metrics that are computed by Garmin's proprietary algorithm, such as heart rate, sleep stage information (i.e., periods and events of light/deep sleep), stress levels, and physical activity levels (including energy expenditure) and step counts). Through the Garmin Health Connect API, Garmin's servers will push the various metrics computed from the raw data to the storage servers at Dartmouth College.

Ecological Momentary Assessment

Participants will be prompted 3 times per day over 12 weeks by the smartphone app to self-report sleep, stress, pain severity, pain interference, pain catastrophizing, craving, withdrawal, substance use risk context, mood, context, substance use, self-regulation, and MOUD adherence (41, 56–59). The EMA prompt delivery times will be randomized within each of the prompt timeframes (e.g., morning, mid-day, end of day). In addition to prompted EMAs, participants will be asked to self-initiate EMA responses if substance use occurred (e.g., opioids, cocaine, or other stimulants). When determining the rate of completion of self-initiated reports of substance use, we will be able to cross-reference responses to the following question asked in the "End of the Day" EMA prompt ("Did you use any drugs at all today without reporting it?") with participant's self-initiated substance use EMA data.

Additionally, participants will be asked to complete a **Momentary Self-Regulation Scale**[1] via EMA. This brief 12-item questionnaire assesses self-regulation on a momentary basis as individuals move through their daily lives. This information will be collected 3 times daily over the 12-week study period by smartphone.

Social Media

Participants will be asked to request and then download their social media data (Facebook, Instagram, or Twitter) to a secure server using a remote desktop application. These three social media platforms provide the functionality for each user through their account setting to download their social media data as an aggregated structured file. After requesting a data download from a social media website, the participant will receive an email notification when the downloadable copy of the data has been created—typically in <48 h from the request. Once the social media data are ready to download, the participant will log into a remote computer located at Dartmouth College by using Microsoft Remote Desktop and will download their social media data to the secure research study computer. After completing the download, the participant will sign out of the remote computer and alert the research team. Images and text postings as well as date/time for each post will be extracted from the downloaded social media data. As noted elsewhere, participation in this part

---

[1] Scherer EA, Kim SJ, Metcalf SA, Sweeney MA, Wu J, Xie H, et al. The development and validation of a momentary self-regulation scale. *JMIR Mental Health* (under review).

of the study is optional; participants can still participate in the study and decline to provide their social media data.

We will parse the JSON/JS files of the downloaded social media data to extract the information of interest, including posting date, text, and corresponding image paths on a local storage. We will aggregate the extracted data into a pickle file composed of different data dictionaries for text, posting dates, and local image paths. It is noteworthy that all the social media data from three variant platforms will be in the same format after processing. We plan to collapse and aggregate the data collected across the three social media platforms to reduce data sparsity and thereby increase the number of study days that are represented in the training and evaluation data sets.

EHR Data

EHR data extraction will include all outpatient and inpatient encounters, medications, procedures, and diagnoses for the 12 months prior to EMA start and the 12-week study period. In addition, we will extract lab results from urine drug screens, patient demographic information, KPNC health plan membership status, and insurance deductible level. We will extract appointment data to determine if visits were canceled or missed. KPNC is also an insurance plan and has claims data on non-KPNC services that were submitted as medical claims for reimbursement.

Clinical and Safety Assessments

Clinical and safety events may be elicited at baseline or spontaneously reported to study staff at any encounter following consent. Safety events suggesting medical or psychiatric deterioration will be brought to the attention of the study clinician for further evaluation and management.

Compensation

Participants will be compensated up to $21 per week for completing EMA surveys, up to an additional $10 per week bonus for completing a minimum of 80% of received EMAs, and up to $14 per week for carrying their smartphone at least 8 h per day and wearing the smartwatch at least 18 h per day. At the end of the 12-week active phase of the study, participants will receive a $50 bonus for either using their personal smartphone or returning a study-provided smartphone, and a $50 bonus for returning the study-provided smartwatch. Finally, participants who consent to the social media portion of the project will receive up to an additional $180. Total possible compensation will be up to $820 over the course of the 12-week active study phase for the digital data collected (i.e., each EMA completed plus EMA bonus, smartphone carry time met and smartwatch wear time met, and social media data download, if applicable). In addition to the earnings and bonuses (described above), an individual who completes a minimum of 80% of received EMA surveys within a given week will qualify for a drawing at the end of that week where the individual could win a $50 prize. Each individual will have an opportunity to participate in up to 12 drawings over the 12 weeks. During the 12-week active study phase, any incentives, bonuses, and/or drawings earned will be uploaded weekly to a reloadable debit card. Study participants will be compensated

$75 for completing the baseline appointments and baseline urine drug screen (via Target gift card), and $100 for completing the 12-week follow-up appointment and follow up urine drug screen (via Target gift card). Total compensation will be up to $995 for participating in all study activities [digital compensation ($820) plus baseline and follow-up appointment compensation ($175)].

Statistical Analyses
Primary and Secondary Outcomes (Endpoints) and Hypotheses

The primary outcomes will include (1) the percentage of days during the 12-week active phase enrolled participants met criteria for wearing the smartwatch and carrying the smartphone; (2) the response rate to EMA prompts during the 12-week active phase; and (3) the percentage of participants who consent to social media data download and sparsity of social media data per participant. We hypothesize that the majority of participants who enroll in the study will wear the smartwatch, carry the smartphone, respond to EMA prompts, and be willing to share their social media data with the research team. We expect the number of participants deemed "non-engagers" will be low.

The secondary outcome measures will be (1) OUD treatment retention (days retained in OUD treatment program) based on EHR data; (2) days covered on MOUD based on EHR and EMA data; and (3) non-prescribed opioid use based on EHR and EMA data. We hypothesize that intensive longitudinal digital data capturing patient context and psychological state will be useful for predicting treatment retention, opioid use events and buprenorphine medication adherence.

Statistical Methods

For our primary feasibility assessment for primary and secondary outcomes, we will generate descriptive statistical summaries of the level of adherence of study participants to the desired protocol (e.g., EMA response rate, smartwatch wear rate and smartphone carry rate).

For our predictive analyses for primary and secondary outcomes, we are interested in measuring and predicting outcomes that may occur repeatedly over a 12-week observational period (e.g., patterns of daily drug use) using digital health technology. In digital health the spatio-temporal granularity of information about an individual is of higher resolution than that obtained through cross-sectional or traditional longitudinal studies (60). We will therefore assess the utility of using data from smartphones, smartwatches, social media, and ecological momentary assessment to predict, explain and detect these outcomes.

Our approach to prediction will include regression methods (e.g., logistic regression), but we will also use various machine-learning approaches for binary classification (e.g., random forest, support vector machines, K-means, gradient boosted trees, neural networks). For each of these classification techniques, we will assess the utility of the various digital data for improving prediction quality.

The study will generate the nested longitudinal data with binary response sequences collected over time. The regression model (logistic regression) will be built to account for the

nested data structures by incorporating both fixed effects and random effects, which would allow us to examine both inter and intra-individual differences. Machine learning models can also be integrated with the random-effects structure as in the mixed-effect models (61). In the cross-validation, the training data will be split into k-folds by patient id. Previous work has shown that whether training data is split by record or by patient can significantly affect model performance (62), with better performance typically being achieved when splitting data by record rather than by participant.

For social media data, we will use deep neural networks for feature extraction and predictive analysis. Specifically, pretrained residual neural network (ResNet) (63) will be used to extract features from images and bidirectional encoder representations from transformers (BERT) (64) models will be used to extract features from text. Using these neural networks, social media images and text can be represented as dense vectors that can be aggregated with the rest of the collected data for predictive analysis. We will also explore classic machine-learning methods (such as random forest, support vector machines, and gradient boosted trees) for social media-based prediction, and compare their results to the performance of deep neural networks. Typical evaluation methods used to assess the prediction quality include area under the receiver operating characteristic curve (AUROC), accuracy, precision, F-score, sensitivity (recall), and specificity. The relative utility of the various data for predicting outcomes will be assessed at two levels - individual features and aggregated features (e.g., Facebook, Twitter, GPS, step, sleep, mood). The contribution of each feature in predicting the outcome variable will be assessed using a model-agnostic machine-learning approach to reverse-engineering algorithms by perturbing model inputs based on game theory, SHapley Additive exPlanations (SHAP) (65, 66).

When missing data are encountered, we will apply domain knowledge to reflect on the probable reasons that the data are missing. Based on our knowledge-based assessment of the nature of the missing data, missing samples will be imputed using appropriate imputation methods (67, 68).

There are at least two approaches to integrating data from the three data sources for use by a single prediction model, depending on whether the prediction models operate in a lower-dimensional "latent/embedding space" or a higher-dimensional "feature space". Deep learning models typically operate in the latent/embedding space, while other machine learning models (e.g., Random Forest, Gradient Boosted Regression Tree) operate in the feature space.

When combining data in the feature space, features must be engineered from both structured and unstructured data. The unstructured social media data, in particular, may require manual or automated annotation in order to generate features. When combining data in the latent space, models that convert both structured and unstructured data into the latent space will be required. These models could be pre-trained on other similar data sources (e.g., BERT for natural language text, pre-trained ResNet model on ImageNet for image data, Activity2Vec for sensor data). We are not aware of pre-trained models for generating embeddings from EMA/survey data.

Yet another approach for "integrating" all 3 data sources is to train separate models using each dataset and an ensemble predictor that combines the predictions from each model to generate a final prediction, e.g., bagging or accuracy-weighted ensemble (69).

We will perform k-fold cross validation (CV) when evaluating the performance of the prediction models. We will do a group k-fold CV where instead of randomly splitting all data into k-folds, we will divide our dataset into k groups such that each participant is assigned to only one group with no overlap between the groups. This is to prevent any data leakage that might happen due to a participant's data being present in the train and test sets.

This is an exploratory pilot study. Therefore, a detailed analysis of statistical power to detect effects was not performed. As we are predicting daily outcomes (e.g., daily medication adherence, daily drug use), the sample size that is potentially available to us is equal to the number of participants multiplied by the number of study days. For example, assuming 60 participants in the study, and a study period of 12 weeks (i.e., 84 days), the analytic sample size would be $60*84 = 5,040$ participant-days. If the observed incidence of non-adherence or drug use is 10%, then we would observe ∼504 non-adherence or drug use events. Results of this study may generate a data set that could be helpful for future researchers to estimate the likely power of predictive models for this patient population, using similar sources of data.

DISCUSSION

In a world that is rapidly embracing digital health approaches to understand and provide resources to support health behavior, this study is distinct in that it will be the first to systematically assess the feasibility and utility of digitally-derived data from EMA, passive sensing and social media, all collected from the same sample of individuals in MOUD treatment. Results from this study may be useful in elucidating novel relations between digital data sources and treatment outcome. It may also inform approaches to enhancing outcomes measurement in clinical trials by allowing for the assessment of dynamic interactions between individuals' daily lives and their MOUD treatment response. It may additionally inform specific digital data collection protocols in the next phase of this line of research, including the need to abbreviate EMA questions to capture those most clinically useful and/or strategies for addressing any privacy or data sharing concerns that may arise among participants. As the opioid epidemic and opioid overdoses surge in the U.S., this novel study and its clinically-relevant implications are timely.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board (IRB) at Kaiser Permanente Northern California (the single IRB overseeing this study). The patients/participants provided their written informed consent to participate in this study. We plan to disseminate findings from this research at scientific meetings and in multiple peer-reviewed publications, including results of analyses assessing the feasibility of using EMA, digital sensing and social media data among adults in outpatient MOUD and analyses assessing the predictive utility of these data sources in predicting MOUD treatment retention, opioid use events, and medication adherence.

AUTHOR CONTRIBUTIONS

LM, CC, C-HC, SH, DK, CL-H, and CS were responsible for developing the study design for this project. LM wrote the initial draft of the protocol paper. AA, CC, MD, MJ-M, HJ, CL-H, and BM were involved in study coordination and operations. GS was the study's Scientific Officer. SA, CL-H, VM, and WW prepared the data sets for analyses. NJ consulted on the statistical analysis plan. C-HC, ZL, VM, and WW conducted the statistical analyses. EH prepared the references. All authors participated in the review and revision process and approved the submission of this version of the protocol paper.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

1. Katz J. *Drug Deaths in American Are Rising Faster Than Ever*. New York, NY: The New York Times (2017).

2. Centers for Disease Control and Prevention (CDC). *Wide-Ranging Online Data for Epidemiologic Research (Wonder)*. Atlanta, GA: National Center for Health Statistics (2018). Available from: https://healthdata.gov/dataset/wide-ranging-online-data-epidemiologic-research-wonder (accessed November 16, 2018).

3. Centers for Disease Control Prevention (CDC), National Center for Health Statistics. Drug overdose deaths in the U.S. top 100,000 annually [Internet]. Atlanta, GA (2021). Available online at: https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2021/20211117.htm#:~:text=Provisional%20data%20from%20CDC's%20National,same%20period%20the%20year%20before (accessed November 17, 2021).

4. Centers for Disease Control Prevention (CDC). *Surveillance for Viral Hepatitis- United States*. Atlanta, GA: U.S. Department of Health and Human Services (2016). Available from: https://www.

cdc.gov/hepatitis/statistics/2016surveillance/index.htm (accessed November 16, 2018).

5. Centers for Disease Control and Prevention (CDC) [Internet]. *New Hepatitis C Infections Nearly Tripled over Five Years.* Atlanta, GA. (2017). Available from: https://www.cdc.gov/nchhstp/newsroom/2017/Hepatitis-Surveillance-Press-Release.html (accessed May 11, 2017).

6. Hartman L, Barnes E, Bachmann L, Schafer K, Lovato J, Files DC. Opiate injection-associated infective endocarditis in the southeastern United States. *Am J Med Sci.* (2016) 352:603–8. doi: 10.1016/j.amjms.2016.08.010

7. Keeshin SW, Feinberg J. Endocarditis as a marker for new epidemics of injection drug use. *Am J Med Sci.* (2016) 352:609–14. doi: 10.1016/j.amjms.2016.10.002

8. Patrick SW, Davis MM, Lehmann CU, Cooper WO. Increasing incidence and geographic distribution of neonatal abstinence syndrome: United States 2009 to 2012. *J Perinatol.* (2015) 35:650–5. doi: 10.1038/jp.2015.36

9. Rhyan CN. *The Potential Societal Benefit of Eliminating Opioid Overdoses, Deaths, and Substance Use Disorders Exceeds $95 Billion Per Year.* Washington, DC: Altarum, Center for Value and Healthcare (2017).

10. Larochelle MR, Bernson D, Land T, Stopka TJ, Wang N, Xuan Z, et al. Medication for opioid use disorder after nonfatal opioid overdose and association with mortality: a cohort study. *Ann Intern Med.* (2018) 169:137–45. doi: 10.7326/M17-3107

11. Koehl JL, Zimmerman DE, Bridgeman PJ. Medications for management of opioid use disorder. *Am J Health Syst Pharm.* (2019) 76:1097–3. doi: 10.1093/ajhp/zxz105

12. Connock M, Juarez-Garcia A, Jowett S, Frew E, Liu Z, Taylor RJ, et al. Methadone and buprenorphine for the management of opioid dependence: a systematic review and economic evaluation. *Health Technol Assess.* (2007) 11:1–171, iii–iv. doi: 10.3310/hta11090

13. Johnson RE, Jaffe JH, Fudala PJ. A controlled trial of buprenorphine treatment for opioid dependence. *JAMA.* (1992) 267:2750–5. doi: 10.1001/jama.1992.03480200058024

14. Ling W, Charuvastra C, Collins JF, Batki S, Brown LS Jr, et al. Buprenorphine maintenance treatment of opiate dependence: a multicenter, randomized clinical trial. *Addiction.* (1998) 93:475–86. doi: 10.1046/j.1360-0443.1998.9344753.x

15. Sordo L, Barrio G, Bravo MJ, Indave BI, Degenhardt L, Wiessing L, et al. Mortality risk during and after opioid substitution treatment: systematic review and meta-analysis of cohort studies. *Bmj.* (2017) 357:j1550. doi: 10.1136/bmj.j1550

16. Martin SA, Chiodo LM, Wilson A. Retention in care as a quality measure for opioid use disorder. *Subst Abus.* (2019) 40:453–8. doi: 10.1080/08897077.2019.1635969

17. Liebschutz JM, Crooks D, Herman D, Anderson B, Tsui J, Meshesha LZ, et al. Buprenorphine treatment for hospitalized, opioid-dependent patients: a randomized clinical trial. *JAMA Intern Med.* (2014) 174:1369–76. doi: 10.1001/jamainternmed.2014.2556

18. Fiellin DA, Barry DT, Sullivan LE, Cutter CJ, Moore BA, O'Connor PG, et al. A randomized trial of cognitive behavioral therapy in primary care-based buprenorphine. *Am J Med.* (2013) 126:74.e11-7. doi: 10.1016/j.amjmed.2012.07.005

19. Gryczynski J, Mitchell SG, Jaffe JH, Kelly SM, Myers CP, O'Grady KE, et al. Retention in methadone and buprenorphine treatment among African Americans. *J Subst Abuse Treat.* (2013) 45:287–92. doi: 10.1016/j.jsat.2013.02.008

20. Klimas J, Hamilton MA, Gorfinkel L, Adam A, Cullen W, Wood E. Retention in opioid agonist treatment: a rapid review and meta-analysis comparing observational studies and randomized controlled trials. *Syst Rev.* (2021) 10:216. doi: 10.1186/s13643-021-01764-9

21. Timko C, Schultz NR, Cucciare MA, Vittorio L, Garrison-Diehn C. Retention in medication-assisted treatment for opiate dependence: a systematic review. *J Addict Dis.* (2016) 35:22–35. doi: 10.1080/10550887.2016.1100960

22. Kampman K, Jarvis M. American society of addiction medicine (asam) national practice guideline for the use of medications in the treatment of addiction involving opioid use. *J Addict Med.* (2015) 9:358–67. doi: 10.1097/ADM.0000000000000166

23. National Quality Forum (NQF). *Behavioral Health 2016–2017: Technical Report.* Washington, DC: Department of Health and Human Services (2017).

24. Fiellin DA, Moore BA, Sullivan LE, Becker WC, Pantalon MV, Chawarski MC, et al. Long-term treatment with buprenorphine/naloxone in primary care: results at 2-5 years. *Am J Addict.* (2008) 17:116–20. doi: 10.1080/10550490701860971

25. Schuman-Olivier Z, Weiss RD, Hoeppner BB, Borodovsky J, Albanese MJ. Emerging adult age status predicts poor buprenorphine treatment retention. *J Subst Abuse Treat.* (2014) 47:202–12. doi: 10.1016/j.jsat.2014.04.006

26. Simon CB, Tsui JI, Merrill JO, Adwell A, Tamru E, Klein JW. Linking patients with buprenorphine treatment in primary care: predictors of engagement. *Drug Alcohol Depend.* (2017) 181:58–62. doi: 10.1016/j.drugalcdep.2017.09.017

27. Weinstein ZM, Kim HW, Cheng DM, Quinn E, Hui D, Labelle CT, et al. Long-term retention in office based opioid treatment with buprenorphine. *J Subst Abuse Treat.* (2017) 74:65–70. doi: 10.1016/j.jsat.2016.12.010

28. DeVito EE, Carroll KM, Sofuoglu M. Toward refinement of our understanding of the fundamental nature of addiction. *Biol Psychiatry.* (2016) 80:172–3. doi: 10.1016/j.biopsych.2016.06.007

29. Bhavnani SP, Narula J, Sengupta PP. Mobile technology and the digitization of healthcare. *Eur Heart J.* (2016) 37:1428–38. doi: 10.1093/eurheartj/ehv770

30. Agrawal R, Prabakaran S. Big data in digital healthcare: lessons learnt and recommendations for general practice. *Heredity.* (2020) 124:525–34. doi: 10.1038/s41437-020-0303-2

31. Lydon-Staley DM, Barnett I, Satterthwaite TD, Bassett DS. Digital phenotyping for psychiatry: accommodating data and theory with network science methodologies. *Curr Opin Biomed Eng.* (2019) 9:8–13. doi: 10.1016/j.cobme.2018.12.003

32. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol.* (2008) 4:1–32. doi: 10.1146/annurev.clinpsy.3.022806.091415

33. Marsch LA. Opportunities and needs in digital phenotyping. *Neuropsychopharmacology.* (2018) 43:1637–8. doi: 10.1038/s41386-018-0051-7

34. Marsch LA. Digital health and addiction. *Curr Opin Syst Biol.* (2020) 20:1–7. doi: 10.1016/j.coisb.2020.07.004

35. Marsch LA. Digital health data-driven approaches to understand human behavior. *Neuropsychopharmacology.* (2021) 46:191–6. doi: 10.1038/s41386-020-0761-5

36. Ferreri F, Bourla A, Mouchabac S, Karila L. E-addictology: an overview of new technologies for assessing and intervening in addictive behaviors. *Front Psychiatry.* (2018) 9:51. doi: 10.3389/fpsyt.2018.00051

37. Hsu M, Ahern DK, Suzuki J. Digital phenotyping to enhance substance use treatment during the Covid-19 pandemic. *JMIR Ment Health.* (2020) 7:e21814. doi: 10.2196/21814

38. Scott CK, Dennis ML, Gustafson DH. Using ecological momentary assessments to predict relapse after adult substance use treatment. *Addict Behav.* (2018) 82:72–8. doi: 10.1016/j.addbeh.2018.02.025

39. Epstein DH, Willner-Reid J, Vahabzadeh M, Mezghanni M, Lin JL, Preston KL. Real-time electronic diary reports of cue exposure and mood in the hours before cocaine and heroin craving and use. *Arch Gen Psychiatry.* (2009) 66:88–94. doi: 10.1001/archgenpsychiatry.2008.509

40. Shiffman S, Waters AJ. Negative affect and smoking lapses: a prospective analysis. *J Consult Clin Psychol.* (2004) 72:192–201. doi: 10.1037/0022-006X.72.2.192

41. Preston KL, Kowalczyk WJ, Phillips KA, Jobes ML, Vahabzadeh M, Lin JL, et al. Before and after: craving, mood, and background stress in the hours surrounding drug use and stressful events in patients with opioid-use disorder. *Psychopharmacology.* (2018) 235:2713–23. doi: 10.1007/s00213-018-4966-9

42. Mun CJ, Finan PH, Epstein DH, Kowalczyk WJ, Agage D, Letzen JE, et al. Craving mediates the association between momentary pain and illicit opioid use during treatment for opioid-use disorder: an ecological momentary assessment study. *Addiction.* (2021) 116:1794–804. doi: 10.1111/add.15344

43. Panlilio LV, Stull SW, Kowalczyk WJ, Phillips KA, Schroeder JR, Bertz JW, et al. Stress, craving and mood as predictors of early dropout from opioid agonist therapy. *Drug Alcohol Depend.* (2019) 202:200–8. doi: 10.1016/j.drugalcdep.2019.05.026

44. Panlilio LV, Stull SW, Bertz JW, Burgess-Hull AJ, Lanza ST, Curtis BL, et al. Beyond abstinence and relapse Ii: momentary relationships between stress,

craving, and lapse within clusters of patients with similar patterns of drug use. *Psychopharmacology.* (2021) 238:1513–29. doi: 10.1007/s00213-021-05782-2

45. Epstein DH, Tyburski M, Kowalczyk WJ, Burgess-Hull AJ, Phillips KA, Curtis BL, et al. Prediction of stress and drug craving ninety minutes in the future with passively collected Gps data. *NPJ Digit Med.* (2020) 3:26. doi: 10.1038/s41746-020-0234-6

46. Sarker H, Tyburski M, Rahman MM, Hovsepian K, Sharmin M, Epstein DH, et al. Finding significant stress episodes in a discontinuous time series of rapidly varying mobile sensor data. *Proc SIGCHI Conf Hum Factor Comput Syst.* (2016) 2016:4489–501. doi: 10.1145/2858036.28 58218

47. Hassanpour S, Tomita N, DeLise T, Crosier B, Marsch LA. Identifying substance use risk based on deep neural networks and instagram social media data. *Neuropsychopharmacology.* (2019) 44:487–94. doi: 10.1038/s41386-018-0247-x

48. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, et al. *The PhenX Toolkit: Substance Abuse and Addiction Core: Tier 1* (2012). Available online at: https://www.phenxtoolkit.org/sub-collections/view/8 (accessed November 19, 2019).

49. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, et al. The phenx toolkit: get the most from your measures. *Am J Epidemiol.* (2011) 174:253–60. doi: 10.1093/aje/kwr193

50. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, et al. *The PhenX Toolkit: Substance Abuse and Addiction Core: Tier 2* (2012). Available online at: https://www.phenxtoolkit.org/sub-collections/view/8 (accessed November 19, 2019).

51. Center for Behavioral Health Statistics and Quality. *2018 National Survey on Drug Use and Health (NSDUH): CAI Specifications for Programming (English Version)*. Rockville, MD: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality (2017).

52. First MB, Williams JBW, Karg RS, Spitzer RL. *Structured Clinical Interview for Dsm-5 Research Version Scid-5-Rv*. Arlington, TX: American Psychiatric Association (2015).

53. Brodey BB, First M, Linthicum J, Haman K, Sasiela JW, Ayer D. Validation of the netscid: an automated web-based adaptive version of the scid. *Compr Psychiatry.* (2016) 66:67–70. doi: 10.1016/j.comppsych.2015. 10.005

54. *Aware.* Open-source Context Instrumentation Framework for Everyone. Available from: https://awareframework.com (accessed July 1, 2019).

55. *Garmin.* Vívosmart® 4. Available from: https://www.garmin.com/en-US/p/605739 (accessed July 1, 2019).

56. Buysse DJ, Reynolds CF, 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res.* (1989) 28:193–213. doi: 10.1016/0165-1781(89)90047-4

57. Freeman LK, Gottfredson NC. Using ecological momentary assessment to assess the temporal relationship between sleep quality and cravings in individuals recovering from substance use disorders. *Addict Behav.* (2018) 83:95–101. doi: 10.1016/j.addbeh.2017. 11.001

58. Kuerbis A, Reid MC, Lake JE, Glasner-Edwards S, Jenkins J, Liao D, et al. Daily factors driving daily substance use and chronic pain among older adults with hiv: an exploratory study using ecological momentary assessment. *Alcohol.* (2019) 77:31–9. doi: 10.1016/j.alcohol.2018.10.003

59. Preston KL, Kowalczyk WJ, Phillips KA, Jobes ML, Vahabzadeh M, Lin JL, et al. Context and craving during stressful events in the daily lives of drug-dependent patients. *Psychopharmacology.* (2017) 234:2631–42. doi: 10.1007/s00213-017-4663-0

60. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. *Nat Biotechnol.* (2015) 33:462–3. doi: 10.1038/nbt.3223

61. Ngufor C, Van Houten H, Caffo BS, Shah ND, McCoy RG. Mixed effect machine learning: a framework for predicting

longitudinal change in hemoglobin A1c. *J Biomed Inform.* (2019) 89:56–67. doi: 10.1016/j.jbi.2018.09.001

62. Saeb S, Lonini L, Jayaraman A, Mohr DC, Kording KP. The need to approximate the use-case in clinical machine learning. *Gigascience.* (2017) 6:1–9. doi: 10.1093/gigascience/gix019

63. He K, Zhang X, Ren S, Sun J, (eds.). Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016). p. 770–8. doi: 10.1109/CVPR.2016.90

64. Devlin J, Chang M, Lee K, Toutanova K, (eds.). Bert: pre-training of deep bidirectional transformers for language understanding. In: *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Minneapolis, MN: Association for Computational Linguistics (2019).

65. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: von Luxburg U, Guyon I, editors. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS '17)*. Red Hook, NY: Curran Associates Inc (2017). p. 4768–77. doi: 10.5555/3295222.3295230

66. Lekkas D, Price G, McFadden J, Jacobson NC. The application of machine learning to online mindfulness intervention data: a primer and empirical example in compliance assessment. *Mindfulness.* (2021) 12:2518–34. doi: 10.1007/s12671-021-01723-4

67. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj.* (2009) 338:b2393. doi: 10.1136/bmj.b2393

68. Ji L, Chow SM, Schermerhorn AC, Jacobson NC, Cummings EM. Handling missing data in the modeling of intensive longitudinal data. *Struct Equ Modeling.* (2018) 25:715–36. doi: 10.1080/10705511.2017.141 7046

69. Wang H, Fan W, Yu PS, Han J, (eds.). Mining concept-drifting data streams using ensemble classifiers. In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery Data Mining (KDD 03)*. Washington, DC (2003).