“I Don’t Like Being Told Just What to Do; I Need to Know Why”: Patient Expectations of Machine Learning-Driven Just-in-Time-Adaptive Interventions (JITAIs) for Prenatal Stress Management

MARA ULLOA, Northwestern University, Computer Science, USA
NEGAR KAMALI, Northwestern University, Computer Science, USA
GLENN J. FERNANDES, Northwestern University, Computer Science, USA
ELIZABETH SOYEMI, Northwestern University Feinberg School of Medicine and Brown University, USA
MIRANDA L. BELTZER, Northwestern University Feinberg School of Medicine, USA
NABIL ALSHURAFA, Northwestern University Feinberg School of Medicine, USA
MAIA JACOBS, Northwestern University, Computer Science and Feinberg School of Medicine, USA

Just-in-time-adaptive interventions (JITAIs) have advanced the ways in which technology supports behavior change, providing tailored interventions at the most appropriate time. With machine learning (ML), these interventions are becoming increasingly sophisticated. However, significant efforts are required to fully understand and address individual expectations of ML-driven JITAIs. In this work, we recruited 20 pregnant people to explore their interest in and expectations of a JITAI designed for prenatal stress prediction and prevention, an area of global public health significance. Using design sessions that combined reflection and design feedback activities, we found that participants were interested in using ML-driven JITAI designs that (1) fit into their daily life by offering flexible engagement, (2) supported them in building a mental model of the underlying JITAI functionality, (3) differentiated between non-adherence and non-compliance, and (4) illustrated apparent reciprocal learning. We use these results to propose designs for future iterations of prenatal stress interventions, particularly those leveraging increasingly complex computational technologies.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**.

Additional Key Words and Phrases: Do, Not, Us, This, Code, Put, the, Correct, Terms, for, Your, Paper

ACM Reference Format:
Mara Ulloa, Negar Kamali, Glenn J. Fernandes, Elizabeth Soyemi, Miranda L. Beltzer, Nabil Alshurafa, and Maia Jacobs. 2018. “I Don’t Like Being Told Just What to Do; I Need to Know Why”: Patient Expectations of Machine Learning-Driven Just-in-Time-Adaptive Interventions (JITAIs) for Prenatal Stress Management. *Proc. ACM Meas. Anal. Comput. Syst.* 37, 4, Article 111 (August 2018), 37 pages. https://doi.org/XXXXXXX.XXXXXXX

Authors’ Contact Information: Mara Ulloa, mara.ulloa@u.northwestern.edu, Northwestern University, Computer Science, Evanston, Illinois, USA; Negar Kamali, negar.kamali@u.northwestern.edu, Northwestern University, Computer Science, Evanston, USA; Glenn J. Fernandes, glenn.fernandes@northwestern.edu, Northwestern University, Computer Science, Chicago, USA; Elizabeth Soyemi, elizabeth_soyemi@brown.edu, Northwestern University Feinberg School of Medicine and Brown University, Chicago, USA; Miranda L. Beltzer, mlbeltzer@gmail.com, Northwestern University Feinberg School of Medicine, Chicago, USA, ; Nabil Alshurafa, nabil@northwestern.edu, Northwestern University Feinberg School of Medicine, Chicago, USA; Maia Jacobs, maia.jacobs@northwestern.edu, Northwestern University, Computer Science and Feinberg School of Medicine, Evanston, USA.

ACM 2476-1249/2018/8-ART111
https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Just-in-time adaptive interventions (JITAIs) provide personalized, timely support tailored to an individual's unique needs and context, offering an adequate amount of support at the precise moment needed [94–97]. Recent evaluations have shown the ability of these technologies to have significant health impacts. JITAIs have been used to help individuals improve mental health, manage chronic conditions, increase physical activity, and reduce substance use [42, 49, 73, 85, 109, 141].

Moreover, implementing ML into JITAs introduces both benefits and challenges. ML enhances JITAIs capabilities to detect behaviors, determine the most effective intervention options for an individual, and identify moments of receptivity, thus offering a highly tailored user experience [6, 68, 88, 102]. New complexities also arise with ML, such as the use of large feature sets to identify or predict complex behaviors [13, 105, 129]. While ML-driven JITAIs hold considerable promise, a lack of insight into end user perceptions and valuation of the technology presents a significant barrier to realizing their expected benefits. Therefore, before developing an ML-driven stress prevention JITAI, in our study we sought to first understand pregnant people's beliefs about how the proposed JITAI support may or may not fit with their daily routines and their expectations from such a system given its anticipated capabilities. Broadly, our research objective was to investigate how such tools may align or disalign with end user goals for stress management during pregnancy, preferences for when and how to use the technology, values that should be embedded into tool design, and concerns about this more intelligent JITAI.

In our work, we recognize that the integration of ML into JITAIs, across its advantages and limitations, raises fundamental design questions about how users want to understand and interact with these interventions and the embedded ML models. As there have been compelling contributions from the Human-Computer Interaction (HCI) research community at the intersection of ML and JITAIs in modeling emotional states of individuals, smartphone overuse, and walking behaviors [60, 102, 105], we undertake this investigation from an HCI perspective, by developing and applying storyboard-based design sessions (Figure 2 in Methods), which draw on participatory design (PD) principles. PD recognizes humans as holders of deep experiential expertise indispensable to system design [8]. JITAI development has outpaced our understanding of how people interpret ML-driven personalization for well-being. Without insight into real-world constraints and user interests, their promise remains uncertain. To advance this relatively uncharted area of JITAI research, we investigate pregnant people's perspectives of a near-future ML-driven JITAI.

We anchor our work in the expertise of pregnant people, investigating their expectations for a near-future ML-driven JITAI. This collaborative groundedness is consequential given that prenatal stress—a global public health crisis—carries profound health risks for both the birthing parent and child. The consequences of prenatal stress manifest in adverse outcomes, including conditions such as postpartum depression, preeclampsia, gestational hypertension, miscarriage, and stillbirth [5, 44, 119, 153]. JITAIs hold great promise for mitigating the detrimental health effects of prenatal stress; however, their ultimate effectiveness hinges on user alignment. If the technology does not resonate with the perceptions and values of pregnant people, its potential for impact remains unrealized.

Prenatal stress is highly dynamic and fluctuates rapidly, making the timely and adaptive support from JITAIs well-suited to address these changes. For this reason, the central contribution of this study is identifying how pregnant people perceive and understand these ML-driven systems. By focusing on people's understanding and concerns of the envisioned JITAI, we provide key insights for designing and developing technology that aligns with human needs and expectations.

Our research is further motivated by recent ML research, which leverages sensors and predictive analytics to look at the possibility of predicting high stress [36, 38, 63, 152]. A notable algorithm for predicting next-day prenatal stress was developed by Ng et al. [101], relying on data collected by electrocardiography (ECG) and inertial measurement unit (IMU) sensors. The recent breakthrough of models capable of predicting next-day stress during pregnancy introduce the opportunity to develop adaptive interventions that aim to prevent, rather than respond to, elevated prenatal stress. These models serve as the motivation for our investigation into pregnant people's perceptions of proactive prenatal stress JITAIs.

Prior scholarship has established the risks of prenatal stress and the promise of ML-based detection, yet the voices of the pregnant people are missing. Our contribution addresses that void by examining how pregnant people regard the anticipated value of these systems.

Specifically, we investigate how pregnant people perceive the required input sources, overall interest in automated detection, and their concerns regarding system integration with daily life, functionality, and algorithmic uncertainty. We organized our study around the following five research questions (RQs):

RQ1: How do pregnant people feel about sharing stress data through a sensor and Ecological Momentary Assessment (EMAs) to guide a stress-prevention intervention?

RQ2: How does the potential to predict next-day stress—using perceived stress and physiological stress—align or disalign with pregnant peoples' needs and goals?

RQ3: How do pregnant people imagine using a JITAI, that could predict next-day stress, in daily life?

RQ4: What questions do pregnant people have about the user of ML to generate next-day stress predictions?

RQ5: What perceived benefits and concerns do pregnant people have towards next-day stress prediction interventions?

In seeking pregnant people's insights of this hypothetical ML-driven JITAI for prenatal stress, we ground our questions on relevant ML properties and on key components of the JITAI framework. Accordingly, each question is centered on its associated interface ML property (e.g., data input or uncertainty of the prediction) and on at least one JITAI component (e.g, tailoring variable or intervention options), which we detail in Method's Table 1.

Recognizing both the complexity of this hypothetical JITAI for prenatal stress and the value of storyboards for decomposing this complexity, we developed five storyboards, one for every research question. Each storyboard illustrated the types of functionality and interactions that are technologically feasible, with the help of two characters: Aria to represent a pregnant user and "LUCA" (Lowering Unwanted Cortisol Activity) to represent the ML-driven JITAI. We used these storyboards to engage participants in providing design feedback on a near-future, yet not currently available, JITAI for prenatal stress. To scaffold the design session, each storyboard guided every individual participant through three modules: an example scenario describing how aspects of the intervention could work within the bounds of existing next-day prenatal stress prediction models ([101]), a reflection activity to engage participants in discussing their expectations of the hypothetical JITAI, and a design feedback activity to facilitate additional future-state ideation. In Method's Figure 2, we broadly illustrate these three modules contained in each of the five storyboards.

Our work offers the following contributions:

(1) Based on design sessions with 20 pregnant people about their prenatal stress and its management, we present findings on patient perceptions and anticipated challenges, which they highlighted as important for engaging with next-day prenatal stress predictions (i.e., the hypothetical ML-driven JITAI).

(2) We translate our findings into design recommendations for ML-driven prenatal stress JITAIs. Specifically, we draw attention to the importance of (a) helping pregnant people build a mental model of JITAI functionality and reinforcement learning (RL) and (b) supporting people's autonomy through flexible engagement and by facilitating non-compliance from pregnant people.

(3) More broadly, we contribute to the growing body of research on patient-facing ML, that intentionality solicits patient expertise.

2 Related Work

2.1 Just-in-time adaptive interventions (JITAIs) for Health

Technological advancements, such as the widespread availability of affordable and accessible smartphones in the early 2000s and the development of sensors for measuring behavior, have proven instrumental in behavioral science. These tools have been particularly effective in advancing digital health efforts for behavior change and personalized support, one emblematic example being JITAIs. The original JITAI architecture by Nahum-Shani et al. [97] is comprised of four critical components supporting both proximal and distal outcomes, which we list, define, and provide examples for below:

Decision points are instances when an intervention decision is triggered. For example, Park et al. [103] propose encouraging physical activity with notifications delivered to the user every three hours (at most four times a day), during waking hours. These instances represent points in time where a decision to send an intervention (or not) could be made.

Intervention options is the set of possible treatments. An example of this is depicted in a JITAI aimed at reducing gambling symptoms with intervention modules, containing an array of activities for the user to benefit from [34].

Tailoring variables are the individual's data used to guide when and which intervention to administer. For instance, Battalio et al. [4] collect sensor and EMA data, including electrical activity from the heart, lung volume and breathing rate, motion, and a variety of EMAs, from smoker adults. The data collected are meant to serve as variables that can determine when to tailor the intervention option, and which option to offer.

Decision rules are the logic (i.e., conditionals to be satisfied with thresholds or ranges) connecting a quantified tailoring variable with intervention options. If a condition is met, then this determines what intervention to offer, when, and to whom. For instance, Spruijt-Metz et al. [121] use the following conditional as a decision rule: if a sedentary, overweight adult has fewer than 150 steps in the previous 40 minutes, then offer this user a randomized intervention, like walking, from an intervention option set.

Proximal outcome is the short term goal of an intervention. Suh et al. [124] explore a JITAI for workplace stress, making mitigation of stress the short-term goal of the intervention.

Distal outcome is the long-term goal, which is targeted through the proximal outcome. In Suh et al. [124]'s study, mitigating stress in the workplace is meant to enhance employee well-being in the long-term.

Following this architecture, the intervention is triggered by contextual information, thus delivering the appropriate dosage of support to the individual for a given health goal at the opportune moment [46]. JITAIs, a specific type of Digital Mental Health Intervention (DMHI), show promise across multiple domains, with notable effectiveness in promoting smoking cessation, providing mental health support, increasing physical activity, and reducing substance use [42, 49, 85, 109, 141]. The prevalence of mobile devices led researchers to focus on merging established behavior intervention philosophies with the possibilities offered by mobile technology [97, 122, 132].

A newer challenge now faces researchers with the rapid integration of ML into the broader healthcare landscape, particularly its role in developing JITAIs, which represent a valuable application within the wider context of patient-facing ML. ML-driven JITAIs are more sophisticated and data-intensive than rule-based JITAIs; for example, they can process information in real time to inform future decision-making [121].

These systems frequently employ ML techniques, such as supervised learning to classify and predict high-risk states (e.g., smoking relapses), and reinforcement learning (RL) to optimize intervention delivery by determining the most appropriate timing, frequency, or content (e.g., suggesting a brief stretch instead of a walk when the user is at work) [48, 73, 108, 121]. ML thus plays a crucial role in enabling JITAIs to manage highly dynamic health states, such as prenatal stress [121, 139]. Given their profound public health implications, ML-driven JITAIs are advancing intervention capabilities, with notable progress in detecting and anticipating moments of vulnerability, receptivity, and relevant tailoring variables [47, 64, 88, 107, 154]. Behavioral lapses that cause individuals to abandon intended health goals can be estimated by analyzing data on physiological changes, cognitive factors, and contextual or environmental triggers. For example, predictive algorithms have been trained and tested on EMA and sensor data to forecast smoking relapses [107] and predict overeating lapses during dietary adherence [47]. Consequently, ML algorithms within JITAIs have the potential to facilitate proactive, preventive care and offer personalized support. Through our work, we argue that to translate this potential into meaningful impact, we must ask individuals if and what they need from such systems and incorporate their viewpoints and attitudes into JITAI design.

While this new generation of JITAIs hold promise for enhancing patient well-being through adaptive capabilities, ML-driven JITAIs face persistent challenges, including substantial data requirements, suboptimal end user engagement, the continued challenge of determining receptivity, and limited interpretability [20, 48, 61, 88, 93, 121, 124]–challenges that underscore the need to actively solicit user input in their design and deployment.

This introduces design complexities that co-design with users can help clarify and navigate, such as accurately modeling the temporal dynamics of a user's state (e.g., the frequency and duration of a patient's symptoms); previous work suggests HCI methods as well-suited to address these challenges in the context of JITAIs [26, 61]. Therefore, as these models continue to improve and become more integrated into JITAIs, it is critical to shift focus from development and involve patients in defining design requirements for this increasingly intelligent technology.

2.2 End User Perceptions of JITAIs

Many JITAI studies have focused on feasibility, engagement, and general acceptability in domains such as smoking cessation and physical activity; for example, JITAI trials for smoking cessation have reported high participant satisfaction and feasibility outcomes and JITAIs for physical activity have shown strong feasibility and acceptability metrics [85, 147]. Systematic reviews of JITAIs focus on defining core components and focus on prospective impact of in-the-moment, adaptive support, but these works typically prioritize the theoretical underpinnings not how users perceive adaptive behavior in anticipation of use [53].

A smaller subset of work has sought to elicit user perspectives on JITAIs, such as a recent focus group study with adult smokers (and smoking cessation professionals) that identified domain-specific preferences; however, they relied on focus groups to explore perspectives not exclusively focused on users, and did not engage non-expert users in co-design, limiting their ability to generate actionable design insights for JITAIs [72]. At the design level, there is a growing call for more involvement of the user during early-stage system ideation [50, 61]. User-centered approaches like the one proposed by Kabir et al. [61] illustrate how early input from end users, elicited through a medium-fidelity self-tracking prototype, can shape tailoring variables and decision rules, providing value during the conceptualization and nascent design stages of a JITAI. This work reinforces our shared commitment to human-centered JITAI design.

Prior studies, where JITAI design was guided by theory and literature, demonstrate the importance of soliciting user feedback. For example, feedback helped show that motivational messaging can increase physical activity, context-aware messages encourage breaks during prolonged sitting, and sensor-based reporting (versus self report) can support smoking cessation while maintaining high user engagement and satisfaction [58, 85, 98, 99, 101, 136, 147]. Together, these contributions illustrate how theoretically grounded systems and personalized delivery strategies can result in meaningful behavioral and engagement outcomes. Building on this foundation, in our study we involve users more purposefully in co-design and earlier in the JITAI design process, focusing not only on evaluating and optimizing implemented systems, but on collaboratively shaping how JITAIs are envisioned, interpreted, and designed prior to development.

User feedback has proven valuable in these contexts. We extend this approach by focusing on early design questions and challenges for JITAIs that intend to embed ML techniques. Our work aims to uncover user concerns about the technology, explore how it can best fit into daily life to support sustained adherence and engagement, and identify opportunities to better align JITAI design with the day-to-day benefit these interventions are intended to provide, before implementation. More specially, our study takes a PD, scenario-based approach to investigate how pregnant people conceptualize and evaluate a hypothetical ML-driven stress JITAI before in-the-wild implementation. By presenting scenarios of ML-driven JITAIs for prenatal stress (Figure 2 in Methods), grounded in the principles of PD, and eliciting user perceptions of fit in daily life, as well as concerns they may have about ML use in this category of JITAI, we aim to address the gap between high-level, expert-driven JITAI frameworks and algorithms and their real-world value, as defined by users.

2.3 Designing Patient-Facing Machine Learning (ML) for Healthcare

ML is increasingly prominent in healthcare, with tools ranging from clinical applications, like image analysis in radiology and treatment selection, to emerging patient-facing interventions for everyday health and well-being management [14, 59, 66, 129–131, 140, 142, 151]. While ML's pattern recognition capabilities drive more accurate diagnoses, streamline clinical workflows, and personalize interventions, its application faces challenges including data quality, technical complexity, interpretability, and ethical considerations [18, 69, 122, 131, 135].

Although ML-driven JITAIs are a relatively new class of adaptive health interventions, design guidance can be drawn from the broader literature on ML systems for healthcare. Prior work has shown that when the design of ML tools in healthcare engages patients, clinicians, and other stakeholders, system usability improves and model development better aligns with individual values and daily life [89, 135]. Participatory approaches can help JITAIs move beyond technical optimization to support meaningful, contextually relevant forms of care, fostering mutual understanding between complex ML processes and the individuals who rely on them [22]. In this context, HCI research plays a pivotal role, as the opacity, complexity, and rapid evolution of ML technologies challenge the definition of stable, patient-centered design requirements [149].

To address these challenges, HCI researchers often employ human-centered or value-sensitive design methods, such as surveys, interviews, or co-design sessions, to engage users throughout development [1, 2, 43, 61]. These methods can be adapted into creative subtypes, including comic boarding and storyboarding, as in our study [29, 69, 133]. Conducting this work in patient-centered contexts requires particular sensitivity, as health outcomes, such as the distal effects of prenatal stress, may be directly at stake. Patient mental models and subjective understandings of their conditions often differ from the assumptions embedded in ML training data [3, 151]. For individuals without ML or technical training, this technology's often inscrutable logic intended to support personal health decisions [75]. Therefore, to engineer more effective and value-aligned ML-driven JITAIs for prenatal stress, we address research questions capturing individual appraisals of these systems and how they wish to interact with such technology.

2.4 Interventions for Prenatal Stress and the Integration of Machine Learning

Prenatal stress is a global health issue posing major health risks for both the pregnant person and their child. In the pregnant person, prenatal stress can lead to pregnancy complications and is a risk factor that predicts perinatal depression and anxiety symptoms [119]. In children, prenatal stress has been linked to slower infant development, childhood diseases, and neurodevelopmental vulnerability to mental health problems [30, 31, 113, 125].

Interventions leveraging cognitive behavioral therapy (CBT), interpersonal therapy, and mindfulness techniques have helped effectively reduce prenatal stress [7, 27, 32, 120]. An evidence-based CBT example is the Mothers

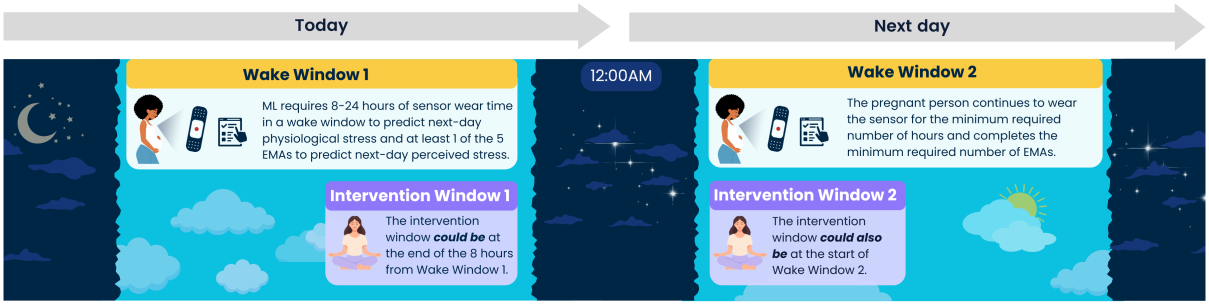

Fig. 1. To predict next-day physiological and perceived stress, a minimum of 8 hours of sensor data is needed and 1/5 of the EMAs need to be answered during the pregnant person's self-identified wake window. These predictions offer the opportunity to provide an intervention at the end of the first wake window or into the start of the next wake window.

and Babies (MB) program [27], with multiple randomized controlled trials (RCTs) showing high effect sizes [28, 126–128, 139]. Though MB alone is efficacious, prior MB trials found individuals experiencing higher stress levels are less likely to receive the full intervention dose, limiting exposure to core MB skills. Recent advances in ML by Ng et al. [101] have facilitated stress monitoring and detection, helping address this limitation. The models developed use EMA and sensor data to predict next-day prenatal perceived and physiological stress (we provide more detail in 3.1). These types of ML advancements enable expanding CBT-based interventions by adding adjunctive interventions that provide in-the-moment support in between CBT training sessions. Combining ML with CBT suggests that continuous stress monitoring can enable JITAIs during pregnancy to reinforce CBT skills, like those taught in MB, in response to real-time stress.

While prior work supports the development of an ML and CBT-driven JITAI for prenatal stress, in our work, we take a step back and explore the opportunities and limitations of integrating ML into JITAIs from the pregnant people's perspective.

3 Methods

To understand whether and how pregnant people would like to engage with a hypothetical stress prediction JITAI, and the underlying ML that would drive such an intervention, we conducted 20 individual design sessions, which took an average of 60 minutes. As the context of our work is motivated by existing stress prediction models developed by Ng et al. [101], in 3.1, we explain how ML-driven stress predictions currently function for pregnant people. In 3.2, we describe our data collection process, covering storyboard creation and rationale. As we will elaborate, we used storyboards to elicit perspectives of potential ML-driven JITAI features. We then describe our participant recruitment process. Finally, in 3.3 we describe our methods for analyzing the data.

3.1 The Development of Next-Day Prenatal Stress Predictions

While our work focuses on perceptions of next-day prenatal stress JITAIs, here we provide some of the background behind the models that have already been created to output next-day stress predictions, highlighting the current state of the technology and the timeliness of our study. Continuous stress monitoring and early prediction of prenatal stress are being facilitated by advances in sensor technology (i.e. adhesive sensor that collects physiological data; see photo in supplemental material) and predictive analytics (i.e., ML). ML research, lead by our co-author, Alshurafa and team Ng et al. [101], demonstrated the feasibility of predicting next-day stress in pregnant people by capturing and analyzing both perceived stress (subjective feeling of being unable to cope) and physiological stress (the body's automatic, biological response of elevated heart rate and cortisol levels to stressors). The perceived stress models were developed to predict next-day stress using EMAs. In parallel, a physiological stress model was developed using electrocardiography (ECG) and inertial measurement unit (IMU) sensor data collected during participants' waking hours (labeled as *"wake window"* in Figure 1). To support model development, participants were asked to wear the sensors for 8 to 24 hours per day. This data was then used to predict the risk of elevated stress over the next 24 hours.

The resulting prenatal stress models successfully predict perceived stress 74.4% of the time and physiological stress 83.6% of the time [101]. This opens the door to a prenatal stress JITAI by using these models to push preventive stress reductions exercises to pregnant people. While the optimal intervention window is not yet known, these predictions present an opportunity to offer preventive intervention at the end of the eight-hour window, and during the following 24 hours, such as at the start of the person's next wake window [63, 79, 101]. However, this remains a hypothetical prenatal stress JITAI, and its true value depends on pregnant people's expectations. Building off Ng et al. [101]'s contribution, which prioritized algorithmic implementation, our work explores key open questions (Table 1) around whether people see value in receiving personalized CBT support tailored by ML to anticipate stressful days.

3.2 Storyboard-based study, Instrument Design and Data Collection, and Participant Recruitment

We begin this section by discussing the value of storyboard-based design sessions, a proven HCI method in helping participants envision the future of technology (3.2.1) [29, 133]. Next, we describe how we designed and applied the storyboards during our design sessions, as the creation and use of the storyboards are mutually informative (3.2.2). Finally, we detail participant recruitment (3.2.3).

3.2.1 Storyboard-based Design Sessions. The existing work on ML-driven JITAIs for prenatal stress focuses primarily on optimizing model performance [101]. However, this alone is insufficient to meet the real-world needs and contexts in which pregnant people would use these tools.

Recognizing the value of storyboard-based design sessions for future-state ideation in driving alignment between technical feasibility and human needs, we created five storyboards based on each of our five research questions (Table 1 and 3.2.2). Storyboards serve as a visual narrative to explore potential solutions early in the design process; they allow participants to collaborate with the researcher, find gaps, and discuss assumptions about the proposed technology before narrowing on design choices, like features, or even developing or deploying the technology [29, 133]. By using storyboards, researchers can facilitate collaborative design sessions that move beyond verbal dialogue, empowering participants to engage with visual representations of envisioned technologies.

Moreover, we used the storyboards to visually guide the participant through a relatable story of a pregnant individual, "Aria". We included a second character, "LUCA" (Lowering Unwanted Cortisol Activity), a text-based agent embodying an ML model, capable of predicting future instances of stress and delivering timely CBT and mindfulness recommendations. In the storyboards we decided to have LUCA interact with Aria via messages, as we've seen a shift toward DMHIs delivered as text messages in recent years [67].

Text messaging is a valuable method for delivering interventions due to its accessibility, especially for individuals who may be hesitant, or simply lack access, to engage with formal mental health services [80]. Prior research has shown that text-based interventions can have comparable efficacy to in-person approaches [78]. As one of the most used forms of interpersonal communication across demographic groups [100], and not requiring internet access [143], there is growing support in the literature for leveraging text messaging and related modalities to meet users where they are in their daily life [33, 41]. As such, we drew significant inspiration from text messaging interventions, recognizing their potential. However, text messaging approaches can often be rudimentary due to engagement challenges. Simple decision trees driving text messaging tend to become repetitive, which have

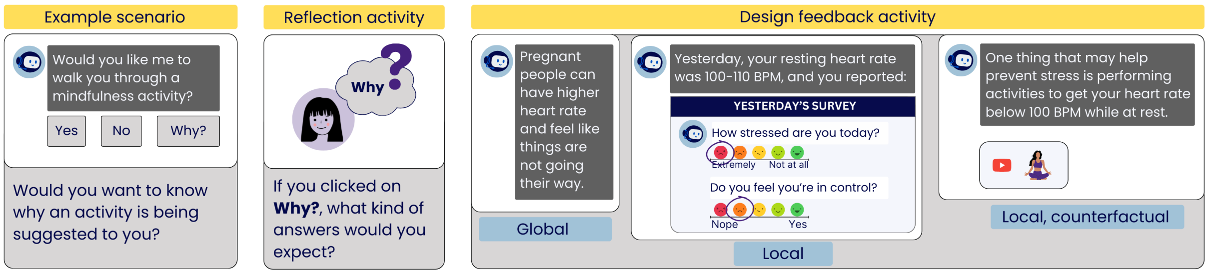

Fig. 2. We designed five storyboards, each focusing on a different property of ML-driven JITAIs. Each storyboard includes three modules: a scenario of potential functionality, a reflection activity, and preliminary designs to foster further reflection and future-state ideation. In the storyboards, "Aria" is a representative end user, and "LUCA" is the ML-JITAI that provides next-day stress alerts and recommendations to reduce prenatal stress.

led to user dissatisfaction and eventual disengagement [65]. In light of this gap, we saw the need for designing LUCA as a more intelligent conversational agent for Aria to engage with, going beyond basic text messaging.

More plainly, we used Aria to help pregnant participants situate themselves and LUCA to facilitate deeper reflection through a potential JITAI, capable of sending intelligent, conversational prompts beyond the limitations of standard text messages. This setup allowed us, for instance, to ask participants to imagine themselves as Aria and consider how they might respond when LUCA delivered messages, such as: *"Imagine you are Aria, and LUCA tells Aria: 'I think today would be a good day for a mindfulness activity [...]', What would you say to LUCA?"* In storyboards described below, we depict and explore in more detail the interactions between Aria and LUCA.

3.2.2 Instrument Design and Data Collection. Here, we further describe our development and application of the storyboards, as these are closely related. Each of the five storyboards included three modules: a scenario, a reflection activity, and a design feedback activity. Figure 2 shows an overview of these three modules, and our supplemental material contains the five storyboards, followed by the list of questions asked to participants. First, we introduced the scenario to the participant, helping describe what was technologically feasible and intervention possibilities. Second, we used reflection to elicit participant comments, questions, concerns, and preferences. Lastly, we offered designs we prepared prior to the design sessions with the goal of soliciting feedback and fostering even further future state ideation with each participant. All participants saw the same five storyboards (i.e., scenarios, reflection activities, and designs for feedback).

We began the sessions by discussing how common stress is for people, especially throughout pregnancy stages, and thanked participants for sharing their personal stress experiences and management strategies with us. Although we recruited and conducted sessions with both pregnant people and those who were up to one week postpartum at the time of the design session, the focus of our study remained on stress experienced during pregnancy. Therefore, when we met with postpartum participants we explicitly asked them to discuss their experiences during pregnancy, rather than the postpartum period.

Given that our work focuses on understanding the perceived value of ML-driven JITAIs, we illustrate in Table 1 how we triangulated each **Research question** (column one) with a storyboard showing an **Interface ML Property** or fit in daily life (column two), and a **JITAI component** (column three). Additionally, in each storyboard described below, we highlight the corresponding RQ number, the Interface ML Property we, and the JITAI component depicted. We based our JITAI components on the original JITAI framework by Nahum-Shani et al. [97]. The following list outlines the design and application of each storyboard, used as our instrument for data collection. Please refer to Table 1 as you read.

Research question (RQ)	Interface ML Property	JITAI component
RQ1: How do pregnant people feel about sharing stress data through a sensor and EMAs to guide a stress-prevention intervention?	Perceived Comfort with **Data Input** Sources	**Decision point:** We asked pregnant people about their comfort with input sources (sensor and answering EMAs), as ongoing data collection through these allows for consideration of multiple possible points for decisions about intervention delivery
RQ2: How does the potential to predict next-day stress—using perceived stress and physiological stress—align or disalign with pregnant peoples' needs and goals?	Perception of **ML Stress Predictions**	**Proximal and distal outcomes:** We asked pregnant people to consider the ways a stress prediction could affect proximal outcomes like next day stress and distal outcomes
RQ3: How do pregnant people imagine using a JITAI, that could predict next-day stress, in daily life?	Perceived **Fit in Daily Life**	**Tailoring variable (receptivity):** We asked pregnant people when in their schedule they would most likely be able to engage with the intervention
RQ4: What questions do pregnant people have about the use of ML to generate next-day stress predictions?	Perception of **Intervention Explanations**	**Intervention options:** We presented pregnant people with the recommendation alongside three types of explanations (local, global, counterfactual) on why a stress intervention is being suggested
RQ5: What perceived benefits and concerns do pregnant people have towards next-day stress prediction interventions?	Perception of **Uncertainty of the Prediction**	**Tailoring variable (frequency):** We presented false positives and false negatives to learn how pregnant people feel about receiving the recommendation more or less often

Table 1. Each of the five storyboards highlighted an interconnected research question, interface ML property (or fit in daily life), and a JITAI component. This method helped us guide inquiry into how the perceptions of pregnant people may align (or not) with a hypothetical ML-driven JITAI for next-day prenatal stress predictions.

(1) **Perceived Comfort with Data Input Sources:**
As we show across the second row of Table 1, **RQ1** in column one guided the design of this storyboard to understand participants comfort and concerns around stress data sources being considred for preventive stress JITAIs. Specifically, we showed participants a sensor and EMA, which were used to collet data for the physiological and perceived stress modeling work by Ng et al. [101].
To capture perceived comfort with **data input** sources (i.e., the Interface ML Property in column two), we presented the participant with a scenario where where Aria uses both the sensor and EMA to capture her stress data. We illustrated Aria using the flexible, band-aid-like, wearable sensor to collect physiological stress data that Ng et al. [101] used, and then depicted Aria completing daily phone-based surveys (i.e., EMAs) to capture perceived stress data. After introducing this initial scenario, we then asked participants to reflect on

the data collection process, probing them on comfort level, data sharing preferences, and any questions and concerns regarding the two types of data collection methods.

Finally, in the design feedback portion, we asked about the frequency at which the participants would be willing to complete surveys and perceptions about wearing the sensor. Understanding pregnant people perceptions of the ongoing data collection via the sensor and EMA is vital, as it is what allows for consideration of multiple **decision points**, which determine a time at which an intervention decision should be made [97]. This corresponds to the JITAI component depicted in column three.

(2) **Perception of ML Stress Prediction:**

We use **RQ2**, presented in the third row of Table 1, along with the information across that row, to guide the creation of this storyboard. We developed this storyboard introducing the **ML stress predictions** (column two), featuring LUCA to the participant. We invited participants to reflect on both the **proximal and distal outcomes** (column three) of the intervention by considering how a just-in-time stress prediction might influence their next-day stress and its longer-term implication [97]. We also asked them to share any questions or concerns about the system's ability to accurately predict stress and deliver appropriate support.

We began the section with a scenario where Aria downloads the LUCA application, which can review Aria's stress data. In this section, LUCA says, *"I reviewed yesterday's sensor and survey data. Based on your data, I predict that completing a mindfulness activity will be helpful. The activity will take a few minutes to complete and can help reduce anxiety and stress levels."*

Next, we asked participants to reflect on whether the predictions made sense to them, what, if any, value they associated with these predictions, and what may factor into their decision-making process to engage with the prediction. We included this as a starting point to encourage participants to consider how predictions are presented to them and what language and tone they prefer.

In the design feedback portion, we focused on whether they require additional interaction with LUCA, such as further questioning, to fully grasp the output predictions' implications.

(3) **Perceived Fit in Daily life:**

In Table 1, row four, we show how this storyboard unpacks **RQ3** to investigate perceived fit of this technology in daily life. This approach helps us assess the potential integration ML interventions into pregnant peoples' everyday routines from their perspective. Understanding **perceived fit in daily life** (column two) is useful for developing **tailoring variables (receptivity)**—variables that guide when the intervention should be offered (column three) [97]. Previous literature also motivated this section, citing receptivity challenges [6] and other obstacles to the integration of mobile health (mHealth) tools in the user's everyday life when misaligned with user priorities [89, 116, 145].

The format of this section was largely informed by Ng et al. [101]. While Ng et al. [101] showed the ability to predict elevated stress using a 24-hour window, as shown in Figure 1, questions remain unanswered regarding when individuals would most like to engage in an intervention to reduce or prevent stress during that time frame. By capturing physiological and perceived stress, using sensors and EMAs, during the first wake window, we may be able offer the pregnant person an intervention:

(a) at the end of the first wake window, once at least eight hours of sensor wear time and at least one EMA have been completed, or

(b) at the beginning of the following wake window.

To capture participant preferences, the scenario we showed participants depicted LUCA chiming in, informing Aria that based on the previous day's data, they would benefit from a mindfulness activity. We use this solution-focused framing, rather than centralizing stress prediction scores, as the suggestion to the individual of future stress (e.g., *"your data suggests you will experience elevated stress levels tomorrow"*) could induce a self-fulfilling prophecy or state of anticipatory anxiety (worrying about the future). We then asked participants to think about their daily routine and evaluate when (morning, daytime, or evening) would be the best time for them to engage with LUCA and possible barriers they foresee.

(4) **Perception of Intervention Explanations**:

The penultimate row in Table 1 shows **RQ4** in column one, which guided our creation of this storyboard. We aimed to capture participant's **perception of intervention explanations** (column two) we designed prior to the design sessions. Following the same structure as the other storyboards, this storyboard focuses on a JITAI component: **intervention options** (column three). To operationalize this component, we presented pregnant participants with stress-management recommendations. Nahum-Shani et al. [97] defines interventions options as the media used to deliver support or the types, sources, or amounts of support. We also accompanied these intervention options with three types of explanations (local, global, and counterfactual), clarifying why an intervention was suggested.

We started the session with the scenario of LUCA suggesting an intervention option, a mindfulness activity, to Aria. The intervention option also allowed Aria to reply to LUCA with *'Yes'* to accept the intervention, *'No'* to reject the intervention, or *'Why?'* to question why LUCA suggested the intervention. We then asked participants if they would be interested in knowing why an activity is being suggested. If so, we asked what kind of answers they would expect to receive from LUCA and to describe it further. After participants described the envisioned explanation they would expect, we showed all participants the same designs we prepared pre-session of three existing explanation types, including global, local, and counterfactual, and asked for their feedback (see supplemental material).

During the design sessions, the global explanation scenario shows Aria receiving an overview of how LUCA makes its recommendations based on data, applicable to the general patient population. We decided to show global and local, as these explanations have been widely used in past work. Global explanations (e.g., feature importance or partial dependence plots) offer a broad understanding of the model to the user; in contrast, local explanations (e.g., SHAP or counterfactual explanations) narrow in on individual predictions [82, 83, 138]. The local explanation scenario showed Aria receiving a personalized explanation for a recommendation from LUCA. The counterfactual scenario illustrated changes to the inputs would have changed a particular outcome. Across the three explanation, we asked participants to share their understanding and feelings on each.

Counterfactuals introduce *"what-if"* scenarios, like what could be changed to get a different result, which help individuals think about causality and control and differ from the descriptive form of global and local explanations [12]. The counterfactual explanation we designed from LUCA to Aria—explicitly shown in supplemental material—is primarily a Better-World Counterfactual (also known as an upward comparison). We selected this type of counterfactual since a Better-World Counterfactual is recommended to help the user understand an ML system and make inferences about its future performance. The counterfactual explanation LUCA shows Aria is also an Additive Counterfactual, as it presents controllable actions that Aria could take [12], aligning with PD's core principle of empowering the user to take control.

(5) **Perception of Uncertainty of the Prediction:**

As shown in the last column of Table 1, **RQ5** informed the creation of this storyboard, focused on understanding pregnant people's **perception of uncertainty of the prediction** (column two). We focused the storyboard around false positive and false negative occurrences, as this can inform JITAI **tailoring variables (frequency)** in column three, guiding decisions around confidence thresholds and the frequency in which recommendations should be provided [97]. More broadly, we aimed to address the inherent uncertainty in ML models, especially related to mental health [24, 84, 151] and capture thoughts of benefits and concerns that participants had towards ML-driven interventions.

We began the design sessions by presenting false positive and false negative scenarios (e.g., *"Imagine LUCA has told you that you are stressed and offers you a mindfulness activity, but you don't feel any stress"*) to the participants. Then, we engaged participants in reflection with a specific question: *"Imagine Aria has been stressed for the last few days. However, LUCA has not noticed the stress and has not sent her a mindfulness activity notification. What would you do if you were in Aria's situation?"* The purpose behind posing this question was threefold. Firstly, it seeks to understand how participants perceive and react to uncertainty and errors in ML recommendations. Second, the participants' responses are instrumental in understanding how users interpret and trust ML-provided information, thereby guiding the development of more transparent ML tools that align with user comprehension and uncertainty tolerance levels. Finally, we asked participants if there were aspects about a tool like LUCA that remained unclear to them or other areas the research team should consider including in the session.

3.2.3 Participant Recruitment. Patients were eligible to partake in our study if they were 18 years and older, pregnant or up to one week postpartum, fluent in English, residing in the United States of America (USA), and not experiencing severe mental health concerns (e.g., bipolar disorder, schizophrenia). With our university's Institutional Review Board approval, we recruited eligible participants during the summer of 2023 from Obstetrics and Gynecology (OB/GYN) inpatient settings within an academic, teaching hospital in the USA. Over a five-week recruitment period, a research team member would routinely screen the Electronic Medical Record (EMR) for hospital patients receiving antepartum care and postpartum patients at the hospital.

To screen participants, eligible antepartum patients were approached at the hospital on the same day of screening, with approval from inpatient clinical staff. Although the study is meant to capture prenatal perspectives, we also recruited participants who were 1–3 days postpartum. These individuals were included because their prenatal experiences were recent and still salient, allowing them to provide meaningful pregnancy-related stress experiences. Additionally, these postpartum patients were only approached if they were clinically stable (i.e. did not experience any adverse delivery or postpartum outcomes such as postpartum hemorrhage) in the hospital during their 1 - 3 days postpartum stay. Including this group allowed for broader participation by those who showed interest in the study's context while maintaining relevance to the prenatal period. All participants, pregnant or recently postpartum, were asked the same questions in the design sessions.

All interested patients were provided with the study flyer, which included a link and QR code to the online screening form. Those eligible proceeded to complete a demographics survey. Both forms were hosted on REDCap (Research Electronic Data Capture), a Health Insurance Portability and Accountability Act (HIPAA)-compliant, secure web-based survey instrument. Participants were contacted via email to schedule their design session with a research team member. All participants chose to meet virtually after being offered the option to meet either virtually or in person. In total, 20 eligible patients (between 27 - 38 years old and currently pregnant or up to one week postpartum) enrolled in the study and participated in the design sessions (Table 2). Participants received a $50 Amazon gift card for their participation.

3.3 Data analysis

We used qualitative thematic analysis [9] to analyze the semi-structured interview data collected as part of our investigation. To initiate the data analysis process, all of the interviews were transcribed via a HIPAA-compliant transcription service company and reviewed by the research team. After this step, we initiated inductive coding of the dataset. Three team members began by independently open-coding a randomly selected 25% of the full dataset. Then, each coder, in turn, presented the research team with initial observations, engaged in open discussion of

	Characteristic	Percentage, n(%)
Race/Ethnicity	White	50%
	Black	30%
	Hispanic/Latino	10%
	Asian	5%
	Unknown	5%
Education	Master's degree or above	45%
	Bachelor's degree	25%
	High school diploma	30%
Income	> $200K	45%
	$100K – $200K	5%
	$50K – $100K	20%
	$25K – $50K	25%
	< $25K	5%
Health Insurance	Private	60%
	Medicaid	20%
	Unknown	20%
Age	23–27	20%
	28–32	35%
	33–38	45%
Marital Status	Married	65%
	Single	15%
	In a committed relationship	15%
	Separated	5%

Table 2. Above, we detail the demographics of the 20 pregnant people who participated in our study.

these observations with the full research team, and examined similarities and differences in our independent analysis during coding comparisons. The first round of coding resulted in substantial overlap related to themes of participant preferences and pain points when using ML for stress predictions. We engaged in five additional rounds of coding with the team. During each round, the research team met to discuss and converge on different perspectives we observed within the data. Through this process, we moved from identifying specific codes to broader categories and patterns within the data. Using the themes established through this iterative process, the first author reviewed all transcripts with the final codebook to validate theme prevalence within the data.

3.4 Positionality Statements

We acknowledge that while our research team includes members from a range of backgrounds, including women, immigrants, ethnic minorities, and individuals across career stages, our composition still represents only a subset of the broader population. As such, we are conscious of how our own lived experiences, positionalities, and disciplinary orientations shaped the study's design, analysis, and interpretation. We approached this work through a constructionist lens, recognizing that knowledge does not simply emerge from the data in a vacuum but is co-constructed through social interaction, both with participants and within the research team.

We developed our thematic analysis through collaborative, interdisciplinary discussion, with coding lead by authors trained in qualitative analysis, and refined in team-wide sessions that included undergraduate, graduate, and senior researchers from ML, HCI, and psychology. These open dialogues served as opportunities to reflect on how our diverse perspectives, informed by a constructionist lens, influenced the ways we interpreted what we were learning from the data. An early-career researcher who was relatively new to JITAI research, for instance, drew attention to how participants expressed a desire to understand the framework itself, rather than simply receiving predictions from LUCA. This perspective led to discussions within the team, particularly with more senior researchers who had extensive experience with JITAI design and modeling. These interactions are directly reflected in our findings, in themes such as participants' need for JITAI framework clarity and participants non-compliance, rejecting a suggested intervention being the end point. Therefore, our thematic analysis not only amplifies participant voices but also captures a balanced integration of diverse perspectives within the research team.

Additionally, ML work previously completed by Alshurafa and team [101] provided an ML lens on the earlier modeling that predates this study; however, our present study is more heavily influenced by HCI principles. In particular, we center the lived experiences of pregnant individuals rather than focusing on optimizing predictive models. Approaching this endeavor through design sessions to empower participants through design thinking further reinforced our constructionist approach, enabling collaborative meaning-making and layered exploration of stress experiences with storyboards, which can be observed in the thoughtful and scaffold nature of the PD method we designed and implemented. Throughout the project, we engaged in reflexive practices challenging each others assumptions and fostering productive disagreement.

4 Findings

Participants detailed several preferences for interacting with ML-driven JITAIs for prenatal stress. Specifically, participants envisioned the use of the intervention in daily life, sharing the following expectations that need to be considered in future designs:

(1) **flexible engagement** to meet the variable demands that arise during pregnancy (4.1), which addresses RQ2, as a tool aligns with a pregnant person's goals only when it meaningfully accommodates the wide range of their day-to-day routines,

(2) support in **building a mental model of JITAI functionality** (4.2), which answers RQ4, as pregnant people have specific questions about system functionality (how does their personal data drives the intervention options?), and RQ5, given that participants concerns can be addressed by providing clear evidence of both short-term outcomes and long-term benefits,

(3) interactions that differentiate between **non-adherence and non-compliance** (4.3), answering RQ3, as individuals envisioned using JITAIs in daily life by engaging with the system's suggestions through real-time feedback and disagreement, and

(4) **reciprocal learning** (4.4), which drove their own willingness to interact with the technology. This answers RQ1, as participants expected their efforts to routinely share stress data via sensors and EMAs should be met with a meaningful benefit—namely, enhanced personalization.

4.1 Engagement Flexibility While Managing Competing Demands During Pregnancy

Participants expressed a strong desire for flexible engagement with ML-driven JITAIs for prenatal stress, because of the high levels of uncertainty associated with this phase of life. This uncertainty was especially pronounced among participants with prior negative pregnancy experiences, who described pregnancy as an inherently unpredictable and complicated period. For example, (P2) shared, *"Overall, I'm excited, but then obviously there's* *also these complications that arise, so it's just not knowing what will come next. It's like this uncertainty. I have no live kids, but I've been pregnant before. I don't know, it's just this uncertainty."* Like (P2), many participants reported experiencing heightened uncertainty during pregnancy. As a result of this uncertainty, they expressed being interested in technology that can anticipate stress, valuing ways to prepare for the unknown and regaining a sense of control.

While participants saw next-day stress predictions as valuable during pregnancy, they revealed needing flexibility to engage with the intervention. Several participants discussed having uncertain daily routines, both due to work and home demands. Accordingly, multiple participants noted they would need to schedule time with LUCA around their work schedule and their children's routines. For instance, (P1) said, *"I do have a [five-year-old] son already, so once I get off work, pick him up from school, it'll just be towards the evening after he had dinner, had his bath, and winds down, then it'll be like my time for me, so then I'll be able to do it."* Similarly, (P19) stated, *"it would be morning before my kids wake up, since I wake up before them, so I can get more stuff done."*

Other participants were grappling with even less certain routines, and required the ability to respond at variable time points. For example, while morning and evening make sense because they are periods of the day for, *"winding up or winding down"* (P8), (P6) added, *"I go through phases. [For example,] I'm in the hospital right now, so I think it just depends on where I am at in life and what's going on."* Importantly, not everyone had a set routine or preferred time to engage LUCA. (P8) shared, *"I'm not a good account of somebody with a normal routine because I very much don't have one. Every day of my week is different; I don't work a normal nine to five job. I'm a freelance voice teacher, so I have a different number of students throughout the day and throughout the week (based on their after school hours or after summer camp). My routine is consistently inconsistent."* The need for flexibility reported by participants directly informs our goal of understanding how an ML-driven JITAI aligns with the needs and goals of pregnant people (RQ2). However, this requirement for user-controlled flexibility highlights a point of friction: it represents a need that stands at odds with the automated, preset logic typical of JITAI design. We explore the implications of balancing system logic with a need for flexibility further in 5.3. Overall, participants perceived next-day stress predictions as being valuable, but communicated the importance of personalized, flexible support that could adapt to different or changing routines.

4.2 Support Pregnant People in Building a Mental Model JITAI Functionality

Using the research questions in Table 1 (see Methods) as a guide, which align with common JITAI components [95], we report on the opportunities participant detailed to help them develop a clearer mental model of how the intervention works.

During the design sessions, **participants shared ways to help them better understand this technology (RQ4) and address concerns about how it works (RQ5)**, in particular across four different components of the JITAI framework (Figure 3). In discussing the design of stress prevention interventions, many participants asked for more information on the sources of the model's predictions and how engaging with the intervention is intended to impact proximal and distal outcomes. Specifically, they wanted to confirm:

(1) their **input data is being valued** and leveraged within the intervention,

(2) how their **input data drives intervention options**,

(3) **evidence supporting the intervention's expected proximal outcome** and,

(4) help **connecting the intervention to long-term benefits**.

Generally, participants expressed that by better understanding how the intervention components were related, they would feel more inclined to accept behavioral recommendations.

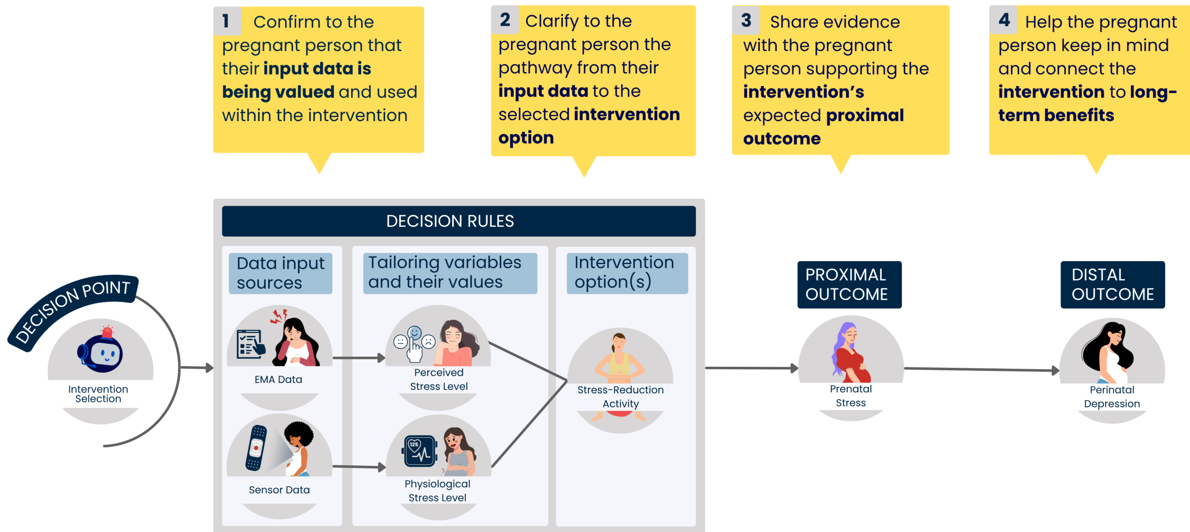

Fig. 3. Above, we illustrate four core components of the JITAI framework that participants wanted to better understand in order to build a coherent mental model of how the system operates. From their perspective, insight into these elements was key for making sense of the tool and addressing skepticism about its intended impact.

4.2.1 ***Confirm input data is being valued:*** Many participants wanted to confirm that the data they shared was being valued and used within the intervention. (P10) expressed that their trust in the intervention would increase if they were assured, and could verify, that the recommendations were based on their own data, *"If I knew that LUCA was taking the direct data from my stress monitor, I would probably trust that it was right and it was working."* (P1) also wanted to know whether their input data was incorporated into the intervention, asking, *"Well, how can you predict that based on today? Where did you get that from?"* They further noted that it would be more logical if the tool explicitly clarified that their data, *"what I reported [via the EMA and sensor]"*, was being used to generate the prediction.

Knowing what types of data (e.g., heart rate, step count and other applicable tailoring variables) are going to be used for selecting an intervention was significant to participants, such as (P16): *"whatever information is going into LUCA's decision-making about why they're recommending it, [the intervention]. If it is the higher heart rate, then being specific about that. If it's step count, if it's based on a conversation that I had with LUCA or whatever, just anything more specific would be helpful."* Participants found it particularly intriguing to observe how the user-supplied data could be applied within the intervention. Pregnant people noted that gaining more insight into this process helped them develop a clearer understanding of how JITAIs function, especially forming a clearer understanding of whether and how individual-level data informs the system's recommendations and decisions.

4.2.2 ***Connect input data to intervention options:*** Beyond wanting to ensure that their input data was reflected within the intervention (as shown above), participants also wanted to know **how** their data guided decision rules and the suggested intervention options. When their data fell outside the typical range, participants were curious about how such deviations influenced the intervention's recommendations and wanted to understand the connection between these changes and the system's guidance. Reflecting during the co-design session, P6 stated, "*So it's giving me information about how to reduce my heart rate below 100 but it doesn't tell me if I need it. Has my heart rate been above 100? Is that why LUCA is worried about me or suggesting this?*" Along the same lines, (P5) expressed a desire to understand the reasoning behind a recommendation, sharing the following expectation, *"I expect the response to be something from the sensor that's just noticing irregular patterns and kind of a recommendation as to how to handle it."* Being able to follow the model's logic helped participants make sense of the system's decisions and increased their confidence in the recommendations: *"knowing why LUCA is deciding what they're deciding,"* would make them feel like LUCA is, *"not being mysterious to me"*, (P16).

Further, participants preferred local explanations and data-backed suggestions, that incorporated their personal data, which was evident when they posed questions such as: "*What in my chart makes him come to the conclusion that I need to do the activity?*" (P7) and "*what is the data that you have that's leading you to suggest this?*" (P4). (P11) described the importance of seeing her input data directly linked to the rules behind the recommendations. She said, *"I would expect to see, 'Well, based on your vitals or whatever you're tracking, your heart rate has been elevated at 80 BPMs, whatever, for this many minutes."* According participants in this study, making this connection visible would clarify how their personal data shapes the system's suggestions and make the design and logic of intervention feel more transparent. Participants wanted to see more than the type of data guiding predictions; they wanted to see how changes in that data drives personalized predictions. Understanding that connection, participants would know more about the relationship between their collected data, ranges in that data related to them specifically or simply that their data influences LUCA's decision-making, and the resulting recommendations.

4.2.3 ***Share evidence with the pregnant person that supports the intervention, explaining the expected proximal outcome:*** Participants envisioned interventions that recommended stress-reduction activities but wanted more comprehensive information about how these activities connect to the intended proximal outcomes. Participants said that this type of clarity would enable understanding of what follows (e.g., benefits they can expect soon after doing the activity) and would allow them to mentally map and set expectations. (P4) requested additional scientific justification for each activity, probing the underlying mechanisms and asking why the activity should produce the outcomes observed by the system, *"Why is this exercise supposed to be effective for what you observed?"* Rather than only questioning its effectiveness, they wanted a deeper explanation of how and why the activity leads to the specific expected impact. Several participants, including (P20), expressed a preference for understanding the reasoning behind the activities, stating,

> *"I don't like being told just what to do;* ***I need to know why.****"* — (P20)

Paticipants explained that understanding the reasoning behind each activity (why it is intended to be effective and how it connects to their experience) would help them make sense of the intervention, clarify its purpose, and grasp the potential short-term benefits of engaging with each activity.

The design session helped ideate on explanations to understand the **'why'** behind and intervention. (P8) shared ideas for how the intervention might provide reasoning behind behavioral recommendation by imagining it could tell personal obstacles and how interventions, like mindfulness, have been shown to address these challenges: *"I am assuming LUCA will probably take the data that I gave him about my stress management or my stress and say, 'Here are things that you posed as things you're going through, things you're dealing with, or things you want to work on, and here's the research that supports why doing a mindfulness activity will help either alleviate that stress or help you manage it' or I think it'll be a data-driven response."* Participants, such as (P18), were also interested in historical evidence that the recommended intervention is effective, imagining that, *"if LUCA's got some proven numbers that it has proven to work [before]"*, then LUCA could and should make that information available to them. Knowing this evidence would make for, "*the strongest argument,*" from the tool for the suggestion being recommended. In sum, participants wanted clear insight into how each intervention option is expected to influence short-term outcomes, whether based on their own data or established scientific research, a clarity that current JITAIs do not provide.

4.2.4 ***Connect the intervention to long-term benefits to demystify the distal outcome:*** Some participants asked for more information on long-term health goals, indicating a need for increasing pregnant people's distal outcome awareness. Participants agreed that casting light on the distal outcome would encourage them to adopt the intervention long-term, particularly if the tool could identify the ultimate benefit for them and communicate, "*here is the end result, the benefits of this, then I might be more inclined.*" (P13) (P6), who disclosed a background in psychology, reflected on how the distal outcomes of JITAIs for managing stress can be difficult to perceive for individuals who are not actively experiencing stress at the moment an intervention is suggested. They noted that without deep knowledge of stress or JITAI mechanisms, the rationale for engaging with certain activities may seem abstract, making it harder for users to appreciate the potential long-term benefits, and said, *"the idea of mindfulness can be helpful for everyone. It does take practice. From a general sense, people might not find it that helpful if they're not feeling stressed in that moment because they might not understand that even if they were feeling stressed yesterday, doing it today could still be helpful."* (P6) continued, elaborating on the importance of making distal outcomes more apparent, and added, *"I guess the question is what's the ultimate goal of tracking something like this?"* They hypothesized that without a clear understanding of the long-term objectives, pregnant people may struggle to see the purpose behind the intervention and remain uncertain about its value. Providing pregnant people with a clearer connection between behavioral recommendations to the long-term benefits, or offering reminders of the long-term benefits throughout use, would enhance understanding of JITAI's ultimate goal.

4.3 Decoupling Non-adherence from Non-compliance

We spoke with participants about the possibility of false positives and false negatives when using ML in stress interventions. Participants consistently stated that in such circumstances they would expect to be given the opportunity to explicitly disagree in real time (RQ3). Through further discussion, **participants requested greater autonomy in the system interactions, and pushed against a recommendation being the end-point.** They sought the option to consider with LUCA when they disagreed with the intervention and as a result didn't feel the need to comply with the suggestion. Participants described that during instances of algorithmic uncertainty, they wanted the following interaction: the ability to input reasons for intentionally choosing not to comply with the intervention's recommendations.

When participants discussed model predictions and the recommendations that followed, they raised the importance of having a way to register disagreement with a model's suggestions. Disagreeing with the system would differentiate their **purposeful decision and reasons for not engaging with the recommendation from unintentional non-adherence**.

Participants consistently drew attention to the value of retaining the final say in decisions, signaling a strong desire for autonomy, which they shared was especially fundamental when using model-based tools. A participant reflected,

> *"It's helpful to have a chatbot potentially come in and help you assess your numbers to make sense of them, but at the end of the day, only I actually know how I'm feeling. So to be able* ***to have a say and advocate for myself would be very necessary****."* – (P10)

Another participant echoed this ideology, making clear that she understood her own lived experiences better than a computational tool like LUCA ever could. She felt strongly that if a stress prediction seemed off or inaccurate, there should be a mechanism within the technology to capture her feedback, which could be part of holding the system accountable: *"Whether you are a person or a chatbot, you have to have accountability. You're wrong... Because, I mean, you, you're a chatbot, I am me, so you can't tell me how I feel."* (P1)

Participants painted a vivid picture of wanting decision support that felt collaborative, not a one-way recommendation imposed on them by a rigid system. Their interest in this idea was so strong that, during the sessions, they began imagining even simple features as meaningful avenues of communication. For instance, (P8) suggested a way to give LUCA real-time feedback: *"a button or something that says, 'I'm experiencing stress' or 'I'm feeling good right now.'"* Such a small feature would allow LUCA to be aware of when the pregnant person disagreed with a recommendation or when the intervention lacked sufficient context. Having this option to push back or provide input would help pregnant people feel empowered to advocate for themselves and preserve their autonomy. Ultimately, retaining the final say, by drawing on their personal knowledge of their own experiences, was a core priority for participants in our study.

4.4 Desire for Reciprocity: Why Pregnant People were Supportive of Sharing Personal Data

In addition to seeking greater autonomy through the ability to disagree with the model and ask why, participants expressed great interest in cultivating a reciprocal dynamic with the system. Broadly, participants described their willingness to share additional data to foster a positive, mutually beneficial relationship with personalized health interventions (RQ1). Participants expressed a strong willingness to share data frequently, even beyond the current input minimum thresholds set for current as sensors and EMAs, viewing this as necessary for maximizing the benefits of AI. They had a notion that these tools require larger datasets and that providing personal information allows the system to deliver more personalized, reciprocal support. This concept of reciprocity emerged from participants such as (P1), who contributed, *"I'm willing to share [my data]. That's not an issue at all. Whenever my feedback is needed, I'm willing to share it."* (P2) echoed the importance of reciprocity with LUCA, noting, *"I know it's an AI, but I also know that what you put into this AI is what it can give out."* Participants anticipated that any additional data they provided, beyond the minimum required, would help the LUCA develop a deeper understanding of their uniqueness, allowing it to generate predictions and recommendations that are tailored to their individual routines and needs. They saw this process as a form of ongoing mutual learning, where their input would directly improve the system's accuracy and relevance to their daily environments. This, in their perspective, would help create a more personalized interaction over time with LUCA. As (P20) explained, *"My understanding of AI, and it could be wrong, is the more you use it or the more you're able to give feedback to it, it could actually improve."* Based on participants current understanding of ML and the role it could play in JITAIs for prenatal stress, participants envisioned that their supplemental data would strengthen the model's learning and enhance its overall accuracy.

Participants anticipated that sharing more data with LUCA would lead to both greater accuracy and more personalized predictions. They explained that by taking on the additional effort of providing LUCA with personal information, the system should become increasingly synched to their individual-level experiences, moving away from broad, generic recommendations. Global or one-size-fits-all explanations were viewed as unhelpful, and participants expressed a clear preference for insights that reflected their unique context. Instead of being artificial, LUCA should become more personal, embedding the supplementary information learned about the participant to inform its interactions with them, like mirroring participant stress managing preferences, as (P15) explained: If, "*I'm taking initiative to,*" provide more information to LUCA, then she would expect that the support provided is not, "*so artificial,*" and not something that would cause her to say, "*Yeah, I could have Googled that...*

> *... I want you to give me something that makes it seem* ***like I'm actually talking to you*** *and that you're really going over this [additional data I'm providing]" – (P15)*

Participants like (P15) wanted to be shown recommendations and explanations that are personalized to them specifically, building off the data they contribute to the system. (P20) elaborated on this thought by saying that without deeper personalization of the tool, "*it feels impersonal, very robotic like something you spit out from your AI algorithm, a generic statement that does not pertain to me, so it does not feel like personalized or customized or that it's taking the input from me into account.*" Reciprocity naturally extends from participant's expectation for autonomy. While participants wanted the ability to disagree with a recommendation (remaining the primary holders of control over decisions), reciprocity captures their willingness to provide additional personal data in return. Participants felt that this reciprocity would strengthen their perception that their contributions play a direct role in the system's evolving personalization. In other words, as reported by participants, having a voice within the preset system logic would motivate people actively contribute, which helps build a cycle where autonomy and collaboration reinforce each other.

4.5 Acknowledgment of User Privacy Concerns

Despite participants' strong willingness to share their data with the model and undertake additional tasks required to do so, as demonstrated in the last section (4.4), this willingness did not completely overshadow their ongoing concerns about data confidentiality and privacy. Participants acknowledged knowing the benefits of contributing personal information to improve personalization, yet they remained mindful of the potential risks, saying that trust in the system and clear safeguards would be essential for them to more feel comfortable sharing sensitive data. (P13) and (P6) discussed disapproval of their data usage by technology. They believed that ML's personalization capabilities needed to be leveraged to do better in this regard than previous technologies. Recounting their dissatisfaction with how social media and digital platforms have collected her data without her feeling like she had received sufficient value for it in return, (P13) told us, "*I'm not getting a lot of value out of this and I'm giving them so much and I'm not getting what I would think of in return and I'm certainly not getting any privacy and there are data breaches all the time. So it's my data and I'm sure I give more data than I even realize.*"

(P6) also stressed the need for confidentiality and privacy when interacting with ML-based health technologies: "*I think there's the privacy thing and what are they doing with the information, so the chatbot's learning your stress and your mood and how is that being used or who can share that data and all that. So I mean I'm probably not as worried about that as some people because I think it's interesting, but I think that lens is always there… Just some sort of promise of confidentiality and privacy and that it's not... Obviously, it has to probably conglomerate your data and use it and look at it to provide the recommendations, but that it's not storing it and sharing it.*"

Several participants added similar thoughts on needing to know who (at the individual and organizational level) has access to the data they are sharing and what other purposes, besides aiding their stress management, it may be used for. Hoping to learn about access to their data, (P4) said, "*You always want to know how many people and organizations have access to this data, and what it's used for, right? How is that interacting with my phone and whatever all it's doing with my data? That's the big thing.*" And since this technology is still fairly new to the general public, participants admitted, "*I don't have enough information to fully judge the usefulness of AI, but my main concern would be privacy and data being shared. I would want to know if it's just being used... How is my data helping this AI technology? And how would it be used to help this AI?*" (P14) These insights from our participants indicate that as AI, specifically ML as covered in this work, becomes more embedded in healthcare, participants recognize and are excited about its potential benefits, but this does not erase the need to give them peace of mind about the use of their medical data.

5 Discussion

While recent advances in sensing and ML have enabled next-day stress prediction, their true potential lies not just in algorithmic performance, but in positioning the needs of pregnant people as the primary driver of the innovation, not an afterthought of the implementation. It is imperative to center on the perspectives, expectations, and lived experiences of pregnant people, ensuring these emergent, complex technologies are designed and developed with and for the people they intend to support.

Through this qualitative study with 20 pregnant people, we discussed an ML-driven JITAI for next-day stress prediction during pregnancy with potential end users. Participants detailed four expectations (as summarized in Table 3 and itemized below) that must be considered in future designs of ML-driven JITAIs for prenatal stress:

- supporting pregnant people in building a mental model of JITAI functionality (RQ4, RQ5),

Research question (RQ) addressed and explanation	Finding	Description of participant expectations
RQ2: A tool aligns with a pregnant person's goals only if it respects the high variation in their actual daily routines	(4.1) **Flexible engagement**	High variation in young families' routines necessitates flexibility, such as the ability to respond even outside the state of vulnerability
RQ4: Users have specific questions about system functionality: how their personal data drives the intervention options **RQ5:** Concerns are managed by providing evidence of outcome efficacy and long-term benefits	(4.2) **Support building a mental model of JITAI functionality**	Request for help better understanding the four different components of the intervention, as detailed in 4.2.1 - 4.2.4 and and in Figure 3
RQ3: Individuals imagine using JITAIs in daily life by interacting with the system's suggestions through real-time disagreement	(4.3) **Decouple non-adherence from non-compliance**	Desire to convey when they explicitly disagreed with the recommendation, rather than just non-adhering (e.g. forgetting)
RQ1: Sharing stress data via sensors and EMAs needs to be met with a meaningful benefit: improved personalization	(4.4) **Desire for reciprocity**	Willingness to share additional data if the tool can reciprocate by improving the accuracy and personalization of their recommendations

Table 3. In the first column, we outline each research question and explain how the corresponding findings (presented in the next column) address it. In the final column, we summarize the four key findings that will be discussed in detail below.

- facilitating reciprocal learning (RQ1),
- flexible engagement (RQ2), and
- decoupling non-adherence from non-compliance (RQ3).

Two of our findings, participant's request to receive support in building mental models of JITAI functionality (4.2) and on facilitating reciprocity (4.4), speak to **pregnant people's desire to gain clearer mental models of JITAIs**. We therefore focus on mental models in our first two discussion sections: (5.1) and (5.2).

In (5.3), we turn from mental models to the concept of autonomy. Here, we specify how **JITAs for pregnant people can better support their autonomy**: illustrating their need for flexible engagement (4.1), which is at odds with the current JITAI framework, and conclude by discussing decoupling non-adherence from non-compliance (4.3) within JITAIs.

Throughout the discussion, our aim is to guide the design of future prenatal stress JITAIs, particularly of those leveraging increasingly complex computational technologies. We propose designs as starting points rather than definitive solutions, as we anticipate and encourage alternative interpretations and approaches to operationalizing our findings.

5.1 Towards Demystifying the JITAI Framework for Pregnant People

In our study, participants frequently asked questions and recommended designs to help them build a mental model of JITAI functionality, specifically seeking to demystify how their input data translated into next-day stress predictions (thereby addressing *RQ4: What questions do pregnant people have about the use of ML to generate next-day stress predictions?*). Based on their feedback (see yellow labels in Figure 3), we identify a space to design JITAIs that make apparent the JITAI framework to laypeople by:

(1) confirming their input data is being valued and leveraged within the intervention,

(2) shaping understanding of how their input data drives intervention options,

(3) providing evidence supporting the intervention's expected proximal outcome, and

(4) helping to connect the intervention to long-term benefits.

Points (3) and (4) answer RQ5: *What perceived benefits and concerns do pregnant people have towards next-day stress prediction interventions?*; addressing concerns by providing the evidence (short-term and long-term benefits) that pregnant people require to move from doubt to relying on JITAIs for stress.

At a high level, the need to build a mental model of the JITAI framework appears conceptually tied to algorithmic transparency—information transfer from a model to a user about its decision-making process—a well-documented principle in human-AI literature [1, 37, 39, 110, 115, 118] and one that has seen significant efforts on helping clinicians interrogate ML systems' algorithms [14, 15, 19, 59, 111, 118, 135, 140, 150]. However, algorithmic transparency is a broad concept, and relying solely on ML research terminology to describe the nuanced phenomena shaped by human perspectives can limit the influence of HCI studies on human-centered ML in healthcare [2]. In the case of ML-JITAIs for prenatal stress, we see a similar need for transparency. Yet, in this context, **participants were more concerned with understanding the broader intervention**, of which the ML model represents only one facet. Below, we review the four JITAI components requiring more clarity, while offering design recommendations for prenatal stress JITAIs.

5.1.1 Clarify the connections between JITAI components.

First, participants wanted to **(1) confirm their input data was being valued and leveraged within the intervention**. This signals a need to help pregnant people visualize their data within the intervention.

To help pregnant people confirm that their data are central to their JITAI experience, we can draw from research on personal health informatics and data sharing, which shows that conceptualizing data use fosters engagement and trust [40, 54, 55, 62, 86, 101, 134]. This could take the form of simple training materials that outline the intervention's purpose, describe what data are collected, and explain how often they are collected. More advanced approaches might visualize user-contributed data (through charts, graphs, or progress indicators) directly within the interface. Linking these visualizations to how the intervention adapts over time could be especially effective. For example, correlating user data with pregnancy stages, as seen in popular pregnancy-related applications, may help users understand how their data is being used. Prior work shows that seeing one's data represented in a system can prompt reflection and motivate continued tracking [55, 101]. Visualizing personal data can reassure users that their input is valued and address engagement challenges with sensors or EMAs, supporting long-term adherence [40, 54, 62, 86, 134]. In the context of pregnancy stress, where people can feel out of control, seeing data validating their emotions can offer a sense of reassurance, helping them cope with changes and reground themselves, as one would with a peer or family member [123]. Our findings suggest not only the value of incorporating personal data, as emphasized in the personal informatics literature, but also validating its use to inform recommendations to the user.

5.1.2 Clarify how personal data drives the intervention. Next, participants wanted to understand why a particular stress reduction exercise was selected for them (i.e., **(2) how their input data drives intervention options**). When shown examples during the design sessions, participants strongly favored local explanations that revealed relevant, personal data. We point to the importance of communicating to the user how their data directly influences the intervention they receive.

Presenting clear, interpretable visuals could demonstrate how personal data informs a type of stress reduction exercise. Local explanations, showing pregnant people why a specific suggestion was made based on their data, would be especially valuable. For example, visualizing how completing certain activities has positively impacted a pregnant person's progress previously could show the expected stress level, when the same or similar exercise was completed, and then the actual stress level following that activity. Such designs align with prior health research indicating that when treatment selection processes are made more understandable to patients, satisfaction with care increases and a more patient-centered paradigm emerges [52, 137]. Similarly, recent visualization research suggests that interpretable visualizations are particularly valuable for patient-facing tools, as they help patients make sense of longitudinal health data [23]. Enabling patients to form a mental model linking their data to intervention options, through visualizations or clear explanations identifying which data segments (e.g., specific days or activity types) led to a recommendation, could for this reason foster both understanding and engagement.

5.1.3 Clarify proximal outcome evidence and distal outcome benefits. Finally, participants asked for **(3) evidence that supports the intervention, explaining the expected proximal outcome** and **(4) clarity on its connection to long-term benefits**. Although, behavior scientists and development teams carefully select outcomes and interventions based on theory and clinical practice [97, 122], these outcomes are not always clear to the end user.

While past work, like JITAI protocols and reviews, outline tailoring variables, decision rules, and adaptive strategies [91, 121, 132], these don't specifically focus on showing the user how and why the intervention is supposed to work. A good example to display the proximal outcome is the use of dashboards in tracking weekly step counts to mitigate sedentary behavior [121]. However, we lack clear examples where the relationship between the distal and proximal outcomes is explicitly reflected in the intervention design, as well as an understanding of whether user mental models align with the technology's underlying logic or not. This lack of clarity lead participants in our study to wonder, ***Why this intervention, and what is the data-backed or proven reason it should affect the proximal outcome right now, and how does success in that proximal outcome contribute to the long-term, distal outcome?*** We argue that communication strategies conveying this rationale are valuable to help users build their own mental models, aligning with Thomas and Bond [132]'s call for, *"more research is clearly needed to understand the psychological underpinnings of response to JITAIs."*

One of the designs participants imagined included providing conceptual models of the intervention components at the time of enrollment. Alternatively, participants wanted to see the intervention impact data–whether derived from literature, user-specific insights, or cohort analysis–detailing short- and long-term effects. For pregnant users, connecting to the distal outcome while tracking their pregnancy (e.g., linking long term impact on the baby) could support engagement [11, 62]. Addressing proximal outcome efficacy and distal outcome benefits upfront aligns with prior literature and supports users in understanding and leveraging the intervention effectively. Research indicates that having clarity of intervention benefits can increase the adoption of digital self-help tools, rivaling human-guided interventions [76, 77, 112].

5.2 Designing for Reciprocity: Trading Continued User Feedback for Deep System Adaptation

Participants in our study expressed a strong willingness to engage longitudinally with JITAIs, driven by the desire to form a reciprocal dynamic with the system (addressing RQ1: *How do pregnant people feel about sharing stress data through a sensor and EMAs to guide a stress-prevention intervention?*). Literature on human–machine reciprocity supports this user desire, describing self-disclosure as a method for fostering deeper collaboration with technology [71], though this reciprocal dynamic faces several obstacles when applying reinforcement learning (RL) to digital interventions.

Importantly, RL has been shown to be helpful for JITAIs, for instance, a contextual bandit RL algorithm was applied to personalize walking suggestions adapting to contextual and dosage factors, but its real-world implementation is challenging [48, 73, 121]. Environmental factors can lead to high noise, data sparsity, and merging multiple data streams that impact the quality and speed of the agent's adaptation. During the time it takes for the RL algorithm's agent to learn and achieve optimal performance, the user may grow impatient and disengage during this latent learning period [73, 144]. At the same time, long-term data sharing, required by RL, can lead to cognitive and time burden which might impact sustained engagement [35]. More plainly, that fact that RL takes time may not be clear to the user. Therefore, communicating when and how RL is taking place, **helping users form a mental model of RL**, is essential to manage user expectations, relying on the behavioral reciprocity individuals are willing to offer.

A major design challenge is to understand how we can best represent or visualize the RL process to JITAI users. Prior work shows that we should consider people often have expertise outside of ML, so the interface should account for this by choosing the right type of representation for lay persons [56, 70, 74, 87]. It will be paramount to design a feature that communicates the RL process, updating the user when the intervention is undergoing optimization by showing that an agent learns in an environment through trial and error, guided by a system of rewards that reinforce desirable actions.

A design could communicate the system undergoing improved performance. For instance, in the context of our study, a prenatal stress JITAI could be shown making its way through a maze collecting positive and negative points. The idea is to show the user prenatal stress JITAI strives tirelessly toward its ultimate goal of performing at its best to help the user, reciprocating the user's ongoing commitment of sharing longitudinal data. Previous work on mental models of RL focused on having university members predict actions and describe the agent's algorithm, which resulted a large cognitive load [3]. Building on their work, it would be beneficial to create more approachable designs for JITAIs that explicitly provide a mental model of RL to lay persons; simpler concepts appear in other interfaces, such as games that depict a character's movement through environments or streak trackers that visually represent progress [92]. RL designs will help ensure JITAIs communicate expectations to the user while the system is adapting, even if the benefits of improved personalization may require patience.

Finally, pregnant people may have distinct motivations, than other populations, that drive their interactions and approach with data-driven systems, such as heightened attention to personal health (during such a formative and vulnerable period) or concerns about fetal outcomes (caring for more then the self) [51, 81]. These factors can influence both their willingness to share personal data and their interest in engaging with ML–based recommendations. Future research should explore how these findings generalize to other populations, as the strong willingness to share data observed in this study may not be as true across groups with varying priorities or risk perceptions. Nonetheless, even among people who are less inclined or available to provide extensive data, exploring RL designs remains valuable. Insights on how a system learns over time, i.e., having a more robust mental model of ML-driven JITAI functionality, could set expectations, enhance engagement, and help improve outcomes across a broader spectrum of JITAI users.

5.3 Fostering Autonomy: Flexible Engagement and Decoupling Non-Adherence from Non-Compliance

5.3.1 Designing for flexible engagement in next-day prenatal stress JITAIs. **Participants in our study liked the idea of a prenatal stress JITAI, if it afforded flexibility in when and how they respond.** To address RQ2 (*How does the potential to predict next-day stress—using perceived stress and physiological stress—align or disalign with pregnant peoples' needs and goals?*), next-day stress predictions align with pregnant people's goals only when the system accounts for real-world, routine variability, allowing for flexible engagement. Participants envisioned flexible engagement in two ways:

(1) aligning the intervention with their routine, which is different for individuals, as some have set routines and others have more unpredictable schedules, and by

(2) providing the ability to dismiss or delay when to respond, because even those with preferred engagement times can't be consistently available to the intervention.

Participants' need for flexibility substantiates and extends earlier work on the importance of humans retaining decision-making autonomy (a psychological need to self-govern), along with JITAI literature on identifying moments of receptivity [16, 21, 25, 57, 68, 88, 90, 96, 96, 114, 124, 146, 151].

Recent work on receptivity provides a promising foundation for designing more adaptive interventions, but predictions of receptivity can be limited, which is why flexible engagement is particularly beneficial for JITAIs. A review published at the end of last year by Nahum-Shani and Murphy [93], for instance, names determining user receptivity (when individuals are most likely to engage with an intervention) as a key challenge in JITAI development. There has been significant work on understanding receptivity by creating models to determine receptivity and investigating what it means for individuals to be receptive [21, 57, 68, 88, 124, 146]. This past work emphasizes the value of understanding individual contextual states to deliver more precisely timed, data-driven interventions. However, even with increasingly sophisticated models of receptivity or growing understanding of receptivity, individual availability cannot be consistently determined, signaling the importance of considering flexible engagement within JITAI design and development.

Complementing receptivity-based adaptation, the flexibility literature showcases a human need for autonomy and control in engaging with interventions. Prior work demonstrates that affording flexible engagement windows (allowing individuals to delay, reschedule, or self-initiate interventions) significantly increases satisfaction and adherence in workplace stress JITAIs [57, 124]. Our findings build on this literature by contributing that flexibility and receptivity are not opposing designs, but complementary avenues for JITAIs. While receptivity modeling enables the system to offer interventions at opportune moments, designing for flexibility respect people's autonomy to decide when those moments truly happen.

For pregnant people and other groups with constrained and fluctuating schedules (such as the care partners studied by Yan et al. [146]), supporting flexible JITAI engagement may be particularly impactful. Simple features, such as a *"snooze"* button to postpone an intervention until after evening routines, or an option to preselect preferred engagement windows (e.g., before children wake up in the morning), could help balance algorithmic receptivity modeling with real-world flexibility. In this way, our proposed designs build on both lines of work, promoting JITAIs that are responsive to individual contexts while respecting their autonomy.

Additionally, given that pregnant people are not always available for JITAI interventions and want more flexibility, we must move beyond current paternalistic approaches to JITAIs, with the goal of clearly communicating the benefits and drawbacks of employing flexibility to postpone an intervention. Understanding the tradeoffs of postponing an interventions allows people to make informed decisions about using flexibility, understanding that only engaging at a more convenient time may be less effective than adhering to the JITAI's *'push'* model. Providing this clarity increases the autonomous flexibility users seek and addresses their desire to form a mental model of how the intervention's components connect, which we presented in 4.2 and discussed in 5.1.

The inclusion of flexibility into the JITAI framework prompts a key question: **does offering flexible engagement preserve the core essence of a *'just-in-time'* intervention as defined in original frameworks by Nahum-Shani et al. [97], and if so, how?** Currently, JITAIs primarily operate under a *'push'* model where pragmatic constraints aim to deliver support when an individual could benefit most. This support is pushed to the user based on system inferences from data (e.g., analyzing sensor and EMA). However, introducing flexibility allows for a *'pull'* mechanism that complements the established *'push'* architecture. In our proposed approach, the JITAI is a hybrid, still operating under the traditional *'push'*, however, this is complemented by user-initiated decisions points that they deem as appropriate time windows outside of the JITAI *'push'* alert. A JITAI micro-randomized trial (MRT) incorporated both *'push'* and *'pull'* options [121]. While this design demonstrated the value of offering user-initiated interactions, it's unclear how this flexibility impacted participants and how this may vary across other contexts. Our analysis indicates a clear need to more critically investigate how flexibility, through an added *'pull'* mechanisms or otherwise, impacts the JITAI framework.

5.3.2 Decoupling non-adherence from non-compliance. In addition to desiring more autonomy through flexible engagement, participants wanted to make it clear to the system when they were purposefully choosing not to engage, not just forgetting about the intervention. This addresses RQ3: *How do pregnant people imagine using a JITAI, that could predict next-day stress, in daily life?* and requires the ability to explain why they may disengage due to rejecting an intervention. Rejecting the intervention can be a result of:

(1) **non-adherence**, due to simply forgetting or being preoccupied, or
(2) **non-compliance**, when purposefully disagreeing and disengaging with the JITAI-proposed intervention; this entails deliberately choosing not to comply with the intervention because the individual disagrees with the suggestion.

The ability to disagree via non-compliance, to consciously choose a path other than the one recommended, is a demonstration of autonomy, a fundamental human need well-documented across mHealth, HCI, and JITAI literature [16, 25, 96, 114].

We have seen this tension between personal choice and technological directive in other work [10, 104, 151]. For example, Burgess et al. [10] discuss patient non-compliance and non-adherence as different behavioral traits, and coined the term *"care frictions"* to describe how healthcare technology can be fundamentally at odds with what people perceive as their true needs. This behavior reflects not a simple neglect of the recommendation but a deliberate resistance to its guidance, emphasizing the need to account for patient non-compliance; as shown by prior work, not using an application or disregarding ML recommendations may represent intuition-based disagreement rather than non-adherence [104, 151]. In sum, participants in this study cleverly conveyed a desire to report their feeling of disagreement, to signal when they are purposefully non-complying versus accidentally disengaging.

Thus, our main recommendation is to integrate designs that decouple non-adherence and non-compliance in JITAI systems. We draw inspiration from work focused on improving clinician-ML interactions, where preserving the human expert's judgment is critical to preventing them from overriding their own expertise in favor of a potentially erroneous ML output [59, 135, 148]. This parallel reinforces the need to preserve the pregnant person's expertise about their own body and stress levels. A design could involve asking the pregnant person if they intend to perform the recommendation, and why or why not. Or, a simple, less intrusive design, directly proposed by participants, would be to display a *"Disagree"* option whenever a recommendation is given. This mechanism would introduce a new, nuanced engagement data point that allows the system to differentiate between purposeful rejection of the guidance (non-compliance) and simply forgetting to use the tool (non-adherence). In the pregnancy context, where individuals are coping with significant physical and social changes [106], affording this autonomy to disagree is especially valuable for handling false positives, such as receiving a stress recommendation when the person feels calm. Capturing this *"Disagree"* signal not only respect their autonomy but can also facilitate learning, for instance, by prompting a message that explains the proactive benefits of practices like mindfulness even when stress is not perceived as present, thereby moving JITAIs beyond current adherence metrics.

6 Limitations and Future Directions

As with all scientific inquiries, our work is shaped by the context in which it was conducted. First, our participants do not fully represent groups historically harmed or excluded from participation in ML-driven health interventions. For instance, high income families have shown higher adoption of digital systems [45, 117], and 45% of participants reported an annual income exceeding $200K. Additionally, the goal of this study was exploratory to gather initial perspectives on ML–driven stress prediction during pregnancy, rather than in the postpartum period.

As a result, our investigation does not capture all relevant viewpoints and our findings may not generalize, for example, the potential value of stress management during the particularly demanding early stages of parenthood, among individuals from lower-income brackets, or lower technology literacy. Moreover, pregnant people may have unique motivations, such as heightened health vigilance or perceived benefits for fetal outcomes that can influence their willingness to share data and their desire to understand or engage with suggestions [51, 81]. Future studies specifically focused on lower income, lower literacy, early parenthood, or JITAIs not intended for pregnancy would offer a valuable and necessary complementary perspectives.

It is important to stress how our work relates to the overturning of Roe V. Wade by the USA's Supreme Court in June 2022. We highlight that the context of our study on pregnancy is impacted by the country's changing legal landscape and, arguably of greater significance, instigating a new set of privacy concerns that demand further scientific inquiry and resolution. For instance, while not all users might yet exhibit sufficient concern about mobile health (mHealth) data privacy post the overturn of Roe V. Wade (e.g., when using period tracking phone applications that can infer pregnancy and abortion) [17], the public's level of concern may evolve alongside the changing political sphere. This perception is also likely to differ state-to-state. The changing relationship between national and local reproductive health policies and rapid development of ML for personal well-being must be prioritized and scrutinized, given the potential impact for it on those most at risk, including birthing persons and their children.

In HCI research, we strive to do more than develop technology by soliciting users expectations and weaving their expertise into digital tool design and development. One example of this is the main finding of our study, which shows that people don't simply want a technologically robust tool that tells them what to do with some level of accuracy; they want help from the tool in building a mental model ML JITAI functionality. By leveraging HCI methods, therefore, we can empower people to design digital interventions that promote their well-being and meets their privacy needs. It is worth noting that these needs may change over time or be dependent on the location of the person.

We also acknowledge that participants who formed part of our study all lived in a state that has preserved birthing people's reproductive decision-making since the 2022 ruling. In that way, participant perspectives of data privacy in our work should not be considered generalizable, considering concerns are likely to vary state by state and other factors like socioeconomic or cultural background. While not the focus of our research, we note that our finding on privacy concerns of ML-driven JITAI for prenatal stress (see subsection 4.5) reaffirms that data privacy must be the central focus of designers, researchers, and developers working in the area of reproductive health.

7 Conclusion

ML is swiftly changing the domain of patient-facing interventions like JITAIs, as it allows monitoring and intervening on more complex behaviors, such as prenatal stress. Much of the existing JITAI literature emphasizes algorithmic and theoretical advances that have enabled this class of digital interventions, while comparatively less attention has been given to directly incorporating patient perspectives and preferences into JITAI design and development. Pregnant people in our study, thus, offer necessary perspectives on how ML should be integrated within health interventions for predicting and managing pregnancy stress. In our research, participants requested the following be considered in future designs:

(1) **support building a mental model of JITAI functionality**,
(2) **reciprocal learning**,
(3) **flexible engagement**, and
(4) interactions that **differentiate between non-adherence and non-compliance**.

We use these discussions to consider how prenatal stress JITAIs can better inform patients of their underlying functionality, a particularly challenging design task as these tools increasingly integrate computational advances in ML. Even though such advances provide an avenue for the development of more personalized and adaptive interventions, they also introduce new challenges related to user understanding of these systems. Accordingly, we propose design strategies can bridge this gap to help patients more meaningfully engage with these interventions and benefit from the value experts envision for these tools.

We do not presume to have introduced an exhaustive list of end-user requirements that should be considered in prenatal stress ML interventions. Rather, this work offers suggestions for creating interpretable, actionable, and engaging interventions that instead of diminishing human control, preserve individual autonomy, support people's well-being management, and encourage technology adoption through reciprocal designs that complement user data collection and engagement efforts.

8 Acknowledgments

This material is based upon work supported by the National Science Foundation (NSF) Graduate Research Fellowship Program (GRFP) under Grant No. DGE-2234667. This research was additionally supported by Underwriters Laboratories (UL) under Grant No. SP0070814 and the Center for Advancing the Safety of Machine Intelligence (CASMI). We extend our sincere thanks to all pregnant people who participated in our study for generously sharing their time, insights, and expertise with our research team. Our gratitude also goes to our colleagues at Northwestern University, including Dr. Benjamin Kaveladze, Chris Romano, Eileen Walsh-Rock, Niharika Gopinath Menon, Dr. Rachel Kornfield, and the Yee Lab (Brittney Renee Williams and Dr. Lynn Yee) for their assistance and guidance during this investigation.

References

[1] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. 2019. Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–13.

[2] Tariq Osman Andersen, Francisco Nunes, Lauren Wilcox, Enrico Coiera, and Yvonne Rogers. 2023. Introduction to the special issue on human-centred AI in healthcare: Challenges appearing in the wild. 12 pages.

[3] Andrew Anderson, Jonathan Dodge, Amrita Sadarangani, Zoe Juozapaitis, Evan Newman, Jed Irvine, Souti Chattopadhyay, Matthew Olson, Alan Fern, and Margaret Burnett. 2020. Mental models of mere mortals with explanations of reinforcement learning. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 10, 2 (2020), 1–37.

[4] Samuel L Battalio, David E Conroy, Walter Dempsey, Peng Liao, Marianne Menictas, Susan Murphy, Inbal Nahum-Shani, Tianchen Qian, Santosh Kumar, and Bonnie Spring. 2021. Sense2Stop: a micro-randomized trial using wearable sensors to optimize a just-in-time-adaptive stress management intervention for smoking relapse prevention. *Contemporary Clinical Trials* 109 (2021), 106534.

[5] Kim S Betts, Gail M Williams, Jakob M Najman, and Rosa Alati. 2015. The relationship between maternal depressive, anxious, and stress symptoms during pregnancy and adult offspring behavioral and emotional problems. *Depression and anxiety* 32, 2 (2015), 82–90.

[6] N Bidargaddi, G Schrader, P Klasnja, J Licinio, and S Murphy. 2020. Designing m-Health interventions for precision mental health support. *Translational psychiatry* 10, 1 (2020), 222.

[7] Sarah E Bledsoe and Nancy K Grote. 2006. Treating depression during pregnancy and the postpartum: a preliminary meta-analysis. *Research on Social Work Practice* 16, 2 (2006), 109–120.

[8] Claus Bossen and Kathleen H Pine. 2023. Batman and Robin in healthcare knowledge work: Human-AI collaboration by clinical documentation integrity specialists. *ACM Transactions on Computer-Human Interaction* 30, 2 (2023), 1–29.

[9] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. doi:10.1191/1478088706qp063oa

[10] Eleanor R Burgess, Elizabeth Kaziunas, and Maia Jacobs. 2022. Care Frictions: A Critical Reframing of Patient Noncompliance in Health Technology Design. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–31.

[11] James Bush, Dilek E Barlow, Jennie Echols, Jasmine Wilkerson, and Katherine Bellevin. 2017. Impact of a mobile health application on user engagement and pregnancy outcomes among Wyoming Medicaid members. *Telemedicine and e-Health* 23, 11 (2017), 891–898.

[12] Ruth MJ Byrne. 2019. Counterfactuals in explainable artificial intelligence (XAI): Evidence from human reasoning.. In *IJCAI*. California, CA, 6276–6282.

[13] Danilo Bzdok and Andreas Meyer-Lindenberg. 2018. Machine learning for precision psychiatry: opportunities and challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 3, 3 (2018), 223–230.

[14] Carrie J. Cai, Emily Reif, Narayan Hegde, Jason Hipp, Been Kim, Daniel Smilkov, Martin Wattenberg, Fernanda Viegas, Greg S. Corrado, Martin C. Stumpe, and Michael Terry. 2019. Human-Centered Tools for Coping with Imperfect Algorithms During Medical Decision-Making. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3290605.3300234

[15] Carrie J. Cai, Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2019. "Hello AI": Uncovering the Onboarding Needs of Medical Practitioners for Human-AI Collaborative Decision-Making. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019), 104:1–104:24. doi:10.1145/3359206

[16] Rafael A. Calvo, Dorian Peters, Daniel Johnson, and Yvonne Rogers. 2014. Autonomy in technology design. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 37–40. doi:10.1145/2559206.2560468

[17] Jiaxun Cao, Hiba Laabadli, Chase H Mathis, Rebecca D Stern, and Pardis Emami-Naeini. 2024. "I Deleted It After the Overturn of Roe v. Wade": Understanding Women's Privacy Concerns Toward Period-Tracking Apps in the Post Roe v. Wade Era. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 813, 22 pages. doi:10.1145/3613904.3642042

[18] Danton S Char, Nigam H Shah, and David Magnus. 2018. Implementing machine learning in health care—addressing ethical challenges. *The New England journal of medicine* 378, 11 (2018), 981.

[19] Baljash Cheema, R. Kannan Mutharasan, Aditya Sharma, Maia Jacobs, Kaleigh Powers, Susan Lehrer, Firas H. Wehbe, Jason Ronald, Lindsay Pifer, Jonathan D. Rich, Kambiz Ghafourian, Anjan Tibrewala, Patrick McCarthy, Yuan Luo, Duc T. Pham, Jane E. Wilcox, and Faraz S. Ahmad. 2022. Augmented Intelligence to Identify Patients With Advanced Heart Failure in an Integrated Health System. *JACC: Advances* 1, 4 (Oct. 2022), 100123. doi:10.1016/j.jacadv.2022.100123 Publisher: American College of Cardiology Foundation.

[20] Haomin Chen, Catalina Gomez, Chien-Ming Huang, and Mathias Unberath. 2022. Explainable medical imaging AI needs human-centered design: guidelines and evidence from a systematic review. *npj Digital Medicine* 5, 1 (Oct. 2022), 156. doi:10.1038/s41746-022-00699-2

[21] Woohyeok Choi, Sangkeun Park, Duyeon Kim, Youn-kyung Lim, and Uichin Lee. 2019. Multi-stage receptivity model for mobile just-in-time health intervention. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 3, 2 (2019), 1–26.

[22] Ron Chrisley. 2020. A Human-Centered Approach to AI Ethics: A Perspective from Cognitive Science. In *The Oxford Handbook of Ethics of AI*, Markus D. Dubber, Frank Pasquale, and Sunit Das (Eds.). Oxford University Press, New York, NY, USA, 0. doi:10.1093/oxfordhb/9780190067397.013.29

[23] Corinna A Christmann, Gregor Zolynski, Alexandra Hoffmann, and Gabriele Bleser. 2017. Effective visualization of long term health data to support behavior change. In *Digital Human Modeling. Applications in Health, Safety, Ergonomics, and Risk Management: Health and Safety: 8th International Conference, DHM 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II 8*. Springer, 237–247.

[24] Michelle Chua, Doyun Kim, Jongmun Choi, Nahyoung G. Lee, Vikram Deshpande, Joseph Schwab, Michael H. Lev, Ramon G. Gonzalez, Michael S. Gee, and Synho Do. 2023. Tackling prediction uncertainty in machine learning for healthcare. *Nature Biomedical Engineering* 7, 6 (June 2023), 711–718. doi:10.1038/s41551-022-00988-x Publisher: Nature Publishing Group.

[25] Simon Coghlan, Jenny Waycott, Amanda Lazar, and Barbara Barbosa Neves. 2021. Dignity, autonomy, and style of company: dimensions older adults consider for robot companions. *Proceedings of the ACM on human-computer interaction* 5, CSCW1 (2021), 1–25.

[26] Daniel DL Coppersmith, Walter Dempsey, Evan M Kleiman, Kate H Bentley, Susan A Murphy, and Matthew K Nock. 2022. Just-in-time adaptive interventions for suicide prevention: Promise, challenges, and future directions. *Psychiatry* 85, 4 (2022), 317–333.

[27] Susan J Curry, Alex H Krist, Douglas K Owens, Michael J Barry, Aaron B Caughey, Karina W Davidson, Chyke A Doubeni, John W Epling, David C Grossman, Alex R Kemper, et al. 2019. Interventions to prevent perinatal depression: US Preventive Services Task Force recommendation statement. *Jama* 321, 6 (2019), 580–587.

[28] S. Darius Tandon, Julie A. Leis, Erin A. Ward, Hannah Snyder, Tamar Mendelson, Deborah F. Perry, Mya Carter, Jaime Hamil, and Huynh-Nhu Le. 2018. Adaptation of an evidence-based postpartum depression intervention: feasibility and acceptability of mothers

and babies 1-on-1. *BMC Pregnancy and Childbirth* 18, 1 (April 2018), 93. doi:10.1186/s12884-018-1726-0

[29] Scott Davidoff, Min Kyung Lee, Anind K Dey, and John Zimmerman. 2007. Rapidly exploring application design through speed dating. In *UbiComp 2007: Ubiquitous Computing: 9th International Conference, UbiComp 2007, Innsbruck, Austria, September 16-19, 2007. Proceedings 9*. Springer, ACM, Innsbruck, Austria, 429–446.

[30] Elysia Poggi Davis, Laura M Glynn, Feizal Waffarn, and Curt A Sandman. 2011. Prenatal maternal stress programs infant stress regulation. *Journal of Child Psychology and Psychiatry* 52, 2 (2011), 119–129.

[31] Roberto de Marco, Giancarlo Pesce, Paolo Girardi, Pierpaolo Marchetti, Marta Rava, Paolo Ricci, and Alessandro Marcon. 2012. Foetal exposure to maternal stressful events increases the risk of having asthma and atopic diseases in childhood. *Pediatric Allergy and Immunology* 23, 8 (2012), 724–729.

[32] Cindy-Lee Dennis and Therese Dowswell. 2013. Psychosocial and psychological interventions for preventing postpartum depression. *Cochrane database of systematic reviews* 2 (2013).

[33] Mallory L Dobias, Robert R Morris, and Jessica L Schleider. 2022. Single-session interventions embedded within Tumblr: acceptability, feasibility, and utility study. *JMIR formative research* 6, 7 (2022), e39004.

[34] NA Dowling, SS Merkouris, CJ Greenwood, GJ Youssef, AC Thomas, CO Hawker, DI Lubman, and SN Rodda. 2025. GamblingLess: In-The-Moment: a mixed-methods acceptability and engagement evaluation of a gambling just-in-time adaptive intervention. *Addiction Science & Clinical Practice* 20, 1 (2025), 80.

[35] Christopher Duckworth, Bethany Cliffe, Brian Pickering, Ben Ainsworth, Alison Blythin, Adam Kirk, Thomas MA Wilkinson, and Michael J Boniface. 2024. Characterising user engagement with mHealth for chronic disease self-management and impact on machine learning performance. *NPJ Digital Medicine* 7, 1 (2024), 66.

[36] Begum Egilmez, Emirhan Poyraz, Wenting Zhou, Gokhan Memik, Peter Dinda, and Nabil Alshurafa. 2017. UStress: Understanding college student subjective stress using wrist-based passive sensing. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. 673–678. doi:10.1109/PERCOMW.2017.7917644

[37] Upol Ehsan, Q Vera Liao, Michael Muller, Mark O Riedl, and Justin D Weisz. 2021. Expanding explainability: Towards social transparency in ai systems. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–19.

[38] Emre Ertin, Nathan Stohs, Santosh Kumar, Andrew Raij, Mustafa al'Absi, and Siddharth Shah. 2011. AutoSense: unobtrusively wearable sensor suite for inferring the onset, causality, and consequences of stress in the field. In *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems* (Seattle, Washington) *(SenSys '11)*. Association for Computing Machinery, New York, NY, USA, 274–287. doi:10.1145/2070942.2070970

[39] Motahhare Eslami, Kristen Vaccaro, Min Kyung Lee, Amit Elazari Bar On, Eric Gilbert, and Karrie Karahalios. 2019. User attitudes towards algorithmic opacity and transparency in online reviewing platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[40] Sarah Faisal, Ann Blandford, and Henry WW Potts. 2013. Making sense of personal health information: challenges for information visualization. *Health informatics journal* 19, 3 (2013), 198–217.

[41] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. 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 mental health* 4, 2 (2017), e7785.

[42] Caroline Free, Rosemary Knight, Steven Robertson, Robyn Whittaker, Phil Edwards, Weiwei Zhou, Anthony Rodgers, John Cairns, Michael G Kenward, and Ian Roberts. 2011. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. *The Lancet* 378, 9785 (2011), 49–55.

[43] Batya Friedman, Peter H Kahn Jr, Alan Borning, and Alina Huldtgren. 2013. Value sensitive design and information systems. In *Early engagement and new technologies: Opening up the laboratory*. Springer, 55–95.

[44] Jason Gardosi, Vichithranie Madurasinghe, Mandy Williams, Asad Malik, and André Francis. 2013. Maternal and fetal risk factors for stillbirth: population based study. *Bmj* 346 (2013).

[45] Nina Goldberg, Christin Herrmann, Paola Di Gion, Volker Hautsch, Klara Hefter, Georg Langebartels, Holger Pfaff, Lena Ansmann, Ute Karbach, and Florian Wurster. 2025. Sociodemographic and Socioeconomic Determinants for the Usage of Digital Patient Portals in Hospitals: Systematic Review and Meta-Analysis on the Digital Divide. *Journal of Medical Internet Research* 27 (2025), e68091.

[46] Stephanie P. Goldstein, Brittney C. Evans, Daniel Flack, Adrienne Juarascio, Stephanie Manasse, Fengqing Zhang, and Evan M. Forman. 2017. Return of the JITAI: Applying a Just-in-Time Adaptive Intervention Framework to the Development of m-Health Solutions for Addictive Behaviors. *International Journal of Behavioral Medicine* 24, 5 (Oct. 2017), 673–682. doi:10.1007/s12529-016-9627-y

[47] Stephanie P Goldstein, Fengqing Zhang, John G Thomas, Meghan L Butryn, James D Herbert, and Evan M Forman. 2018. Application of machine learning to predict dietary lapses during weight loss. *Journal of diabetes science and technology* 12, 5 (2018), 1045–1052.

[48] Suat Gönül, Tuncay Namlı, Ahmet Coşar, and İsmail Hakkı Toroslu. 2021. A reinforcement learning based algorithm for personalization of digital, just-in-time, adaptive interventions. *Artificial Intelligence in Medicine* 115 (2021), 102062.

[49] David H Gustafson, Fiona M McTavish, Ming-Yuan Chih, Amy K Atwood, Roberta A Johnson, Michael G Boyle, Michael S Levy, Hilary Driscoll, Steven M Chisholm, Lisa Dillenburg, et al. 2014. A smartphone application to support recovery from alcoholism: a randomized

clinical trial. *JAMA psychiatry* 71, 5 (2014), 566–572.

[50] Chenxu Hao, Susanne Uusitalo, Caroline Figueroa, Quirine TS Smit, Michael Strange, Wen-Tseng Chang, MI Ribeiro, Vanita Kouomogne Nana, Myrthe L Tielman, and Maaike HT de Boer. 2025. A human-centered perspective on research challenges for hybrid human artificial intelligence in lifestyle and behavior change support. *Frontiers in Digital Health* 7 (2025), 1544185.

[51] Haijing Hao, Yang W Lee, Marianne Sharko, Qilu Li, and Yiye Zhang. 2025. Privacy Concerns Versus Personalized Health Content—Pregnant Individuals' Willingness to Share Personal Health Information on Social Media: Survey Study. *JMIR Formative Research* 9 (2025), e60862.

[52] Yuexing Hao, Zeyu Liu, Robert N. Riter, and Saleh Kalantari. 2024. Advancing Patient-Centered Shared Decision-Making with AI Systems for Older Adult Cancer Patients. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 437, 20 pages. doi:10.1145/3613904.3642353

[53] Wendy Hardeman, Julie Houghton, Kathleen Lane, Andy Jones, and Felix Naughton. 2019. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. *International Journal of Behavioral Nutrition and Physical Activity* 16, 1 (2019), 31.

[54] Mariann Hardey. 2019. On the body of the consumer: performance-seeking with wearables and health and fitness apps. *Sociology of Health & Illness* 41, 6 (2019), 991–1004.

[55] Ralf Hartmann, Christian Sander, Noah Lorenz, Daniel Böttger, Ulrich Hegerl, et al. 2019. Utilization of patient-generated data collected through mobile devices: insights from a survey on attitudes toward mobile self-monitoring and self-management apps for depression. *JMIR mental health* 6, 4 (2019), e11671.

[56] Andrea Hartzler and Wanda Pratt. 2011. Managing the personal side of health: how patient expertise differs from the expertise of clinicians. *Journal of medical Internet research* 13, 3 (2011), e62.

[57] Esther Howe, Jina Suh, Mehrab Bin Morshed, Daniel McDuff, Kael Rowan, Javier Hernandez, Marah Ihab Abdin, Gonzalo Ramos, Tracy Tran, and Mary P Czerwinski. 2022. Design of digital workplace stress-reduction intervention systems: Effects of intervention type and timing. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–16.

[58] Tasnim Ismail and Dena Al Thani. 2022. Design and evaluation of a just-in-time adaptive intervention (JITAI) to reduce sedentary behavior at work: experimental study. *JMIR formative research* 6, 1 (2022), e34309.

[59] Maia Jacobs, Jeffrey He, Melanie F. Pradier, Barbara Lam, Andrew C Ahn, Thomas H McCoy, Roy H Perlis, Finale Doshi-Velez, and Krzysztof Z Gajos. 2021. Designing AI for trust and collaboration in time-constrained medical decisions: a sociotechnical lens. In *Proceedings of the 2021 chi conference on human factors in computing systems*. Association for Computing Machinery, Yokohama, Japan, 1–14.

[60] Sooyeon Jeong and Cynthia Lynn Breazeal. 2016. Improving Smartphone Users' Affect and Wellbeing with Personalized Positive Psychology Interventions. In *Proceedings of the Fourth International Conference on Human Agent Interaction* (Biopolis, Singapore) *(HAI '16)*. Association for Computing Machinery, New York, NY, USA, 131–137. doi:10.1145/2974804.2974831

[61] Kazi Sinthia Kabir, Stacey A Kenfield, Erin L Van Blarigan, June M Chan, and Jason Wiese. 2022. Ask the users: a case study of leveraging user-centered design for designing Just-in-Time Adaptive Interventions (JITAIs). *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–21.

[62] Taewan Kim, Haesoo Kim, Ha Yeon Lee, Hwarang Goh, Shakhboz Abdigapporov, Mingon Jeong, Hyunsung Cho, Kyungsik Han, Youngtae Noh, Sung-Ju Lee, et al. 2022. Prediction for retrospection: Integrating algorithmic stress prediction into personal informatics systems for college students' mental health. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–20.

[63] Zachary D. King, JUDITH MOSKOWITZ, BEGUM EGILMEZ, SHIBO ZHANG, LIDA ZHANG, MICHAEL BASS, JOHN ROGERS, ROOZBEH GHAFFARI, LAURIE WAKSCHLAG, and NABIL ALSHURAFA. 2019. micro-Stress EMA: A Passive Sensing Framework for Detecting in-the-wild Stress in Pregnant Mothers. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 3, 3 (Sept. 2019), 91. doi:10.1145/3351249

[64] Zachary D King, Han Yu, Thomas Vaessen, Inez Myin-Germeys, and Akane Sano. 2024. Investigating Receptivity and Affect Using Machine Learning: Ecological Momentary Assessment and Wearable Sensing Study. *JMIR mHealth and uHealth* 12, 1 (2024), e46347.

[65] Rachel Kornfield, Jonah Meyerhoff, Hannah Studd, Ananya Bhattacharjee, Joseph Jay Williams, Madhu Reddy, and David C Mohr. 2022. Meeting users where they are: user-centered design of an automated text messaging tool to support the mental health of young adults. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.

[66] Rachel Kornfield, Prathusha K Sarma, Dhavan V Shah, Fiona McTavish, Gina Landucci, Klaren Pe-Romashko, and David H Gustafson. 2018. Detecting recovery problems just in time: Application of automated linguistic analysis and supervised machine learning to an online substance abuse forum. *Journal of medical Internet research* 20, 6 (2018), e10136.

[67] Rachel Kornfield, Caitlin A Stamatis, Ananya Bhattacharjee, Bei Pang, Theresa Nguyen, Joseph J Williams, Harsh Kumar, Sarah Popowski, Miranda Beltzer, Christopher J Karr, et al. 2023. A text messaging intervention to support the mental health of young adults: user engagement and feedback from a field trial of an intervention prototype. *Internet Interventions* 34 (2023), 100667.

[68] Florian Künzler, Varun Mishra, Jan-Niklas Kramer, David Kotz, Elgar Fleisch, and Tobias Kowatsch. 2020. Exploring the State-of-Receptivity for mHealth Interventions. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 4, Article 140 (sep 2020), 27 pages. doi:10.1145/3369805

[69] Tzu-Sheng Kuo, Hong Shen, Jisoo Geum, Nev Jones, Jason I Hong, Haiyi Zhu, and Kenneth Holstein. 2023. Understanding frontline workers' and unhoused individuals' perspectives on ai used in homeless services. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.

[70] Himabindu Lakkaraju, Jon Kleinberg, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2017. The selective labels problem: Evaluating algorithmic predictions in the presence of unobservables. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. 275–284.

[71] Yi-Chieh Lee, Naomi Yamashita, Yun Huang, and Wai Fu. 2020. "I Hear You, I Feel You": Encouraging Deep Self-disclosure through a Chatbot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3313831.3376175

[72] Corinna Leppin, Tosan Okpako, Claire Garnett, Olga Perski, and Jamie Brown. 2025. Exploring perspectives on digital smoking cessation just-in-time adaptive interventions: A focus group study with adult smokers and smoking cessation professionals. *PLOS Digital Health* 4, 5 (2025), e0000705.

[73] Peng Liao, Kristjan Greenewald, Predrag Klasnja, and Susan Murphy. 2020. Personalized HeartSteps: A Reinforcement Learning Algorithm for Optimizing Physical Activity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 1, Article 18 (mar 2020), 22 pages. doi:10.1145/3381007

[74] Q Vera Liao, Daniel Gruen, and Sarah Miller. 2020. Questioning the AI: informing design practices for explainable AI user experiences. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–15.

[75] Q. Vera Liao, Hariharan Subramonyam, Jennifer Wang, and Jennifer Wortman Vaughan. 2023. Designerly Understanding: Information Needs for Model Transparency to Support Design Ideation for AI-Powered User Experience. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 9, 21 pages. doi:10.1145/3544548.3580652

[76] Kjersti R Lillevoll, Hans Christian B Vangberg, Kathleen M Griffiths, Knut Waterloo, and Martin R Eisemann. 2014. Uptake and adherence of a self-directed internet-based mental health intervention with tailored e-mail reminders in senior high schools in Norway. *BMC psychiatry* 14 (2014), 1–11.

[77] Jiaxi Lin, Bianca Faust, David Daniel Ebert, Lena Krämer, and Harald Baumeister. 2018. A web-based acceptance-facilitating intervention for identifying patients' acceptance, uptake, and adherence of internet-and mobile-based pain interventions: randomized controlled trial. *Journal of medical Internet research* 20, 8 (2018), e244.

[78] Jake Linardon, Pim Cuijpers, Per Carlbring, Mariel Messer, and Matthew Fuller-Tyszkiewicz. 2019. The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. *World Psychiatry* 18, 3 (2019), 325–336.

[79] Yuhao Liu, Limei Tian, Milan S Raj, Matthew Cotton, Yinji Ma, Siyi Ma, Bryan McGrane, Arjun V Pendharkar, Nader Dahaleh, Lloyd Olson, et al. 2018. Intraoperative monitoring of neuromuscular function with soft, skin-mounted wireless devices. *NPJ digital medicine* 1, 1 (2018), 19.

[80] Shaun Liverpool, Catarina Pinheiro Mota, Célia MD Sales, Anja Čuš, Sara Carletto, Camellia Hancheva, Sónia Sousa, Sonia Conejo Cerón, Patricia Moreno-Peral, Giada Pietrabissa, et al. 2020. Engaging children and young people in digital mental health interventions: systematic review of modes of delivery, facilitators, and barriers. *Journal of medical Internet research* 22, 6 (2020), e16317.

[81] Julianne Louie, Tara Mukund, Chau Vu, Daniel A Epstein, and Alexandra Papoutsaki. 2025. Understanding Temporality of Reflection in Personal Informatics through Baby Tracking. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–18.

[82] Scott M. Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M. Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. 2020. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nature machine intelligence* 2, 1 (Jan. 2020), 56–67. doi:10.1038/s42256-019-0138-9

[83] Scott M Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. *Advances in neural information processing systems* 30 (2017).

[84] Ujunwa Madububambachu, Augustine Ukpebor, and Urenna Ihezue. 2024. Machine learning techniques to predict mental health diagnoses: A systematic literature review. *Clinical Practice and Epidemiology in Mental Health: CP & EMH* 20 (2024), e17450179315688.

[85] Jacqueline Louise Mair, Lawrence D Hayes, Amy K Campbell, Duncan S Buchan, Chris Easton, and Nicholas Sculthorpe. 2022. A personalized smartphone-delivered just-in-time adaptive intervention (JitaBug) to increase physical activity in older adults: mixed methods feasibility study. *JMIR formative research* 6, 4 (2022), e34662.

[86] Lena Mamykina, Elizabeth D. Mynatt, and David R. Kaufman. 2006. Investigating health management practices of individuals with diabetes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 927–936. doi:10.1145/1124772.1124910

[87] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence* 267 (2019), 1–38.

[88] Varun Mishra, Florian Künzler, Jan-Niklas Kramer, Elgar Fleisch, Tobias Kowatsch, and David Kotz. 2021. Detecting Receptivity for mHealth Interventions in the Natural Environment. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 2, Article 74 (jun 2021), 24 pages. doi:10.1145/3463492

[89] David C Mohr, Ken R Weingardt, Madhu Reddy, and Stephen M Schueller. 2017. Three problems with current digital mental health research... and three things we can do about them. *Psychiatric services* 68, 5 (2017), 427–429.

[90] Arlen C Moller, Edward L Deci, and Richard M Ryan. 2006. Choice and ego-depletion: The moderating role of autonomy. *Personality and social psychology bulletin* 32, 8 (2006), 1024–1036.

[91] Andre Matthias Müller, Ann Blandford, and Lucy Yardley. 2017. The conceptualization of a Just-In-Time Adaptive Intervention (JITAI) for the reduction of sedentary behavior in older adults. *Mhealth* 3 (2017), 37.

[92] Elizabeth L Murnane, Xin Jiang, Anna Kong, Michelle Park, Weili Shi, Connor Soohoo, Luke Vink, Iris Xia, Xin Yu, John Yang-Sammataro, et al. 2020. Designing ambient narrative-based interfaces to reflect and motivate physical activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[93] Inbal Nahum-Shani and Susan A Murphy. 2025. Just-in-Time Adaptive Interventions: Where Are We Now and What Is Next? *Annual Review of Psychology* 77 (2025).

[94] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. [n. d.]. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support | Annals of Behavioral Medicine | Oxford Academic. https://academic.oup.com/abm/article/52/6/446/4733473

[95] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. 2017. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine* 52, 6 (Dec. 2017), 446–462. doi:10.1007/s12160-016-9830-8

[96] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. 2018. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine* 0, 0 (2018), 1–17.

[97] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. 2018. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine* 52, 6 (May 2018), 446–462. doi:10.1007/s12160-016-9830-8

[98] Felix Naughton, Chloë Brown, Juliet High, Caitlin Notley, Cecilia Mascolo, Tim Coleman, Garry Barton, Lee Shepstone, Stephen Sutton, A Toby Prevost, et al. 2021. Randomised controlled trial of a just-in-time adaptive intervention (JITAI) smoking cessation smartphone app: the Quit Sense feasibility trial protocol. *BMJ open* 11, 4 (2021), e048204.

[99] Felix Naughton, Sarah Hopewell, Neal Lathia, Rik Schalbroeck, Chloë Brown, Cecilia Mascolo, Andy McEwen, Stephen Sutton, et al. 2016. A context-sensing mobile phone app (Q sense) for smoking cessation: a mixed-methods study. *JMIR mHealth and uHealth* 4, 3 (2016), e5787.

[100] Frank Newport. 2014. The new era of communication among Americans. *Gallup News* (2014).

[101] Ada Ng, Boyang Wei, Jayalakshmi Jain, Erin A Ward, S Darius Tandon, Judith T Moskowitz, Sheila Krogh-Jespersen, Lauren S Wakschlag, and Nabil Alshurafa. 2022. Predicting the Next-Day Perceived and Physiological Stress of Pregnant Women by Using Machine Learning and Explainability: Algorithm Development and Validation. *JMIR mHealth and uHealth* 10, 8 (Aug. 2022), e33850. doi:10.2196/33850

[102] Adiba Orzikulova, Han Xiao, Zhipeng Li, Yukang Yan, Yuntao Wang, Yuanchun Shi, Marzyeh Ghassemi, Sung-Ju Lee, Anind K Dey, and Xuhai Xu. 2024. Time2Stop: Adaptive and Explainable Human-AI Loop for Smartphone Overuse Intervention. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–20. doi:10.1145/3613904.3642747

[103] Junghwan Park, Meelim Kim, Mohamed El Mistiri, Rachael Kha, Sarasij Banerjee, Lisa Gotzian, Guillaume Chevance, Daniel E Rivera, Predrag Klasnja, and Eric Hekler. 2023. Advancing understanding of just-in-time states for supporting physical activity (Project JustWalk JITAI): protocol for a System ID study of just-in-time adaptive interventions. *JMIR Research Protocols* 12, 1 (2023), e52161.

[104] Joonyoung Park and Uichin Lee. 2023. Understanding disengagement in just-in-time mobile health interventions. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 2 (2023), 1–27.

[105] Junghwan Park, Gregory J Norman, Predrag Klasnja, Daniel E Rivera, Eric Hekler, et al. 2023. Development and validation of multivariable prediction algorithms to estimate future walking behavior in adults: retrospective cohort study. *JMIR mHealth and uHealth* 11, 1 (2023), e44296.

[106] Kelsey Perrykkad, Rebecca O'Neill, and Sharna D Jamadar. 2024. Sense of self in first-time pregnancy. *npj Women's Health* 2, 1 (2024), 12.

[107] Olga Perski, Dimitra Kale, Corinna Leppin, Tosan Okpako, David Simons, Stephanie P Goldstein, Eric Hekler, and Jamie Brown. 2024. Supervised machine learning to predict smoking lapses from Ecological Momentary Assessments and sensor data: Implications for just-in-time adaptive intervention development. *PLOS Digital Health* 3, 8 (2024), e0000594.

[108] Olga Perski, Kezhi Li, Nikolas Pontikos, David Simons, Stephanie P Goldstein, Felix Naughton, and Jamie Brown. 2023. Classification of lapses in smokers attempting to stop: A supervised machine learning approach using data from a popular smoking cessation smartphone app. *Nicotine and Tobacco Research* 25, 7 (2023), 1330–1339.

[109] Tianchen Qian, Ashley E Walton, Linda M Collins, Predrag Klasnja, Stephanie T Lanza, Inbal Nahum-Shani, Mashfiqui Rabbi, Michael A Russell, Maureen A Walton, Hyesun Yoo, et al. 2022. The microrandomized trial for developing digital interventions: Experimental

design and data analysis considerations. *Psychological methods* 27, 5 (2022), 874.

[110] Emilee Rader, Kelley Cotter, and Janghee Cho. 2018. Explanations as mechanisms for supporting algorithmic transparency. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.

[111] Niroop Channa Rajashekar, Yeo Eun Shin, Yuan Pu, Sunny Chung, Kisung You, Mauro Giuffre, Colleen E Chan, Theo Saarinen, Allen Hsiao, Jasjeet Sekhon, Ambrose H Wong, Leigh V Evans, Rene F. Kizilcec, Loren Laine, Terika Mccall, and Dennis Shung. 2024. Human-Algorithmic Interaction Using a Large Language Model-Augmented Artificial Intelligence Clinical Decision Support System. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 442, 20 pages. doi:10.1145/3613904.3642024

[112] Talia R Cohen Rodrigues, Thomas Reijnders, Linda D Breeman, Veronica R Janssen, Roderik A Kraaijenhagen, Douwe E Atsma, and Andrea WM Evers. 2024. Use Intention and User Expectations of Human-Supported and Self-Help eHealth Interventions: Internet-Based Randomized Controlled Trial. *JMIR Formative Research* 8, 1 (2024), e38803.

[113] R Jeanne Ruiz and Kay C Avant. 2005. Effects of maternal prenatal stress on infant outcomes: a synthesis of the literature. *Advances in Nursing Science* 28, 4 (2005), 345–355.

[114] Lars Sandman and Christian Munthe. 2009. Shared decision-making and patient autonomy. *Theoretical medicine and bioethics* 30 (2009), 289–310.

[115] Bianca G. S. Schor, Chris Norval, Ellen Charlesworth, and Jatinder Singh. 2024. Mind The Gap: Designers and Standards on Algorithmic System Transparency for Users. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 476, 16 pages. doi:10.1145/3613904.3642531

[116] Jaydip. Sen. 2012. Ubiquitous Computing: Applications, Challenges and Future Trends. In *Embedded Systems and Wireless Technology*. CRC Press, UK, online, 23–62. doi:10.1201/b12298-5

[117] Sachin D Shah. 2024. From Innovation to Inclusion—Tackling Digital Equity Needs in Health Care. *JAMA Network Open* 7, 11 (2024), e2445334–e2445334.

[118] Aubrey A Shick, Christina M Webber, Nooshin Kiarashi, Jessica P Weinberg, Aneesh Deoras, Nicholas Petrick, Anindita Saha, and Matthew C Diamond. 2024. Transparency of artificial intelligence/machine learning-enabled medical devices. *NPJ Digital Medicine* 7, 1 (2024), 21.

[119] Shefaly Shorey, Cornelia Yin Ing Chee, Esperanza Debby Ng, Yiong Huak Chan, Wilson Wai San Tam, and Yap Seng Chong. 2018. Prevalence and incidence of postpartum depression among healthy mothers: A systematic review and meta-analysis. *Journal of psychiatric research* 104 (2018), 235–248.

[120] Laura E Sockol. 2015. A systematic review of the efficacy of cognitive behavioral therapy for treating and preventing perinatal depression. *Journal of affective disorders* 177 (2015), 7–21.

[121] Donna Spruijt-Metz, Benjamin M Marlin, Misha Pavel, Daniel E Rivera, Eric Hekler, Steven De La Torre, Mohamed El Mistiri, Natalie M Golaszweski, Cynthia Li, Rebecca Braga De Braganca, et al. 2022. Advancing behavioral intervention and theory development for mobile health: the HeartSteps II protocol. *International journal of environmental research and public health* 19, 4 (2022), 2267.

[122] Donna Spruijt-Metz and Wendy Nilsen. 2014. Dynamic Models of Behavior for Just-in-Time Adaptive Interventions. *IEEE Pervasive Computing* 13, 3 (July 2014), 13–17. doi:10.1109/MPRV.2014.46 Conference Name: IEEE Pervasive Computing.

[123] Aleksandra A Staneva, Fiona Bogossian, and Anja Wittkowski. 2015. The experience of psychological distress, depression, and anxiety during pregnancy: A meta-synthesis of qualitative research. *Midwifery* 31, 6 (2015), 563–573.

[124] Jina Suh, Esther Howe, Robert Lewis, Javier Hernandez, Koustuv Saha, Tim Althoff, and Mary Czerwinski. 2024. Toward tailoring just-in-time adaptive intervention systems for workplace stress reduction: exploratory analysis of intervention implementation. *JMIR mental health* 11 (2024), e48974.

[125] Elizabeth J Susman, Karen H Schmeelk, Brenda K Worrall, Douglas A Granger, Angelo Ponirakis, and George P Chrousos. 1999. Corticotropin-releasing hormone and cortisol: Longitudinal associations with depression and antisocial behavior in pregnant adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry* 38, 4 (1999), 460–467.

[126] S. Darius Tandon, Julie A. Leis, Tamar Mendelson, Deborah F. Perry, and Karen Kemp. 2014. Six-Month Outcomes from a Randomized Controlled Trial to Prevent Perinatal Depression in Low-Income Home Visiting Clients. *Maternal and Child Health Journal* 18, 4 (May 2014), 873–881. doi:10.1007/s10995-013-1313-y

[127] S. Darius Tandon, Deborah F. Perry, Tamar Mendelson, Karen Kemp, and Julie A. Leis. 2011. Preventing perinatal depression in low-income home visiting clients: A randomized controlled trial. *Journal of Consulting and Clinical Psychology* 79, 5 (2011), 707–712. doi:10.1037/a0024895 Place: US Publisher: American Psychological Association.

[128] S Darius Tandon, Erin A Ward, Jaime L Hamil, Cindy Jimenez, and Mya Carter. 2018. Perinatal depression prevention through home visitation: a cluster randomized trial of mothers and babies 1-on-1. *Journal of Behavioral Medicine* 41 (2018), 641–652.

[129] Anja Thieme, Danielle Belgrave, and Gavin Doherty. 2020. Machine learning in mental health: A systematic review of the HCI literature to support the development of effective and implementable ML systems. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 5 (2020), 1–53.

[130] Anja Thieme, Maryann Hanratty, Maria Lyons, Jorge Palacios, Rita Faia Marques, Cecily Morrison, and Gavin Doherty. 2023. Designing Human-centered AI for Mental Health: Developing Clinically Relevant Applications for Online CBT Treatment. *ACM Trans. Comput.-Hum. Interact.* 30, 2, Article 27 (March 2023), 50 pages. doi:10.1145/3564752

[131] Anja Thieme, Abhijith Rajamohan, Benjamin Cooper, Heather Groombridge, Robert Simister, Barney Wong, Nicholas Woznitza, Mark A Pinnock, Maria T Wetscherek, Cecily Morrison, et al. 2025. Challenges for responsible AI design and workflow integration in healthcare: a case study of automatic feeding tube qualification in radiology. *ACM Transactions on Computer-Human Interaction* 32, 4 (2025), 1–61.

[132] J. Graham Thomas and Dale S. Bond. 2015. Behavioral response to a just-in-time adaptive intervention (JITAI) to reduce sedentary behavior in obese adults: Implications for JITAI optimization. *Health Psychology* 34, Suppl (Dec. 2015), 1261–1267. doi:10.1037/hea0000304

[133] Khai N Truong, Gillian R Hayes, and Gregory D Abowd. 2006. Storyboarding: an empirical determination of best practices and effective guidelines. In *Proceedings of the 6th conference on Designing Interactive systems*. 12–21.

[134] Meghan Reading Turchioe, Annie Myers, Samuel Isaac, Dawon Baik, Lisa V Grossman, Jessica S Ancker, and Ruth Masterson Creber. 2019. A systematic review of patient-facing visualizations of personal health data. *Applied clinical informatics* 10, 04 (2019), 751–770.

[135] Mara Ulloa, Blaine Rothrock, Faraz S. Ahmad, and Maia Jacobs. 2022. Invisible clinical labor driving the successful integration of AI in healthcare. *Frontiers in Computer Science* 4 (Dec. 2022), 0. doi:10.3389/fcomp.2022.1045704 Publisher: Frontiers.

[136] Saskia Van Dantzig, Murtaza Bulut, Martijn Krans, Anouk Van Der Lans, and Boris De Ruyter. 2018. Enhancing physical activity through context-aware coaching. In *Proceedings of the 12th EAI international conference on pervasive computing Technologies for Healthcare*. 187–190.

[137] Julia CM van Weert, Barbara C Van Munster, Remco Sanders, René Spijker, Lotty Hooft, and Jesse Jansen. 2016. Decision aids to help older people make health decisions: a systematic review and meta-analysis. *BMC medical informatics and decision making* 16 (2016), 1–20.

[138] Sahil Verma, John Dickerson, and Keegan Hines. 2021. Counterfactual Explanations for Machine Learning: Challenges Revisited. *X* 0 (2021), 0.

[139] Lauren S. Wakschlag, Darius Tandon, Sheila Krogh-Jespersen, Amelie Petitclerc, Ashley Nielsen, Rhoozbeh Ghaffari, Leena Mithal, Michael Bass, Erin Ward, Jonathan Berken, Elveena Fareedi, Peter Cummings, Karen Mestan, Elizabeth S. Norton, William Grobman, John Rogers, Judith Moskowitz, and Nabil Alshurafa. 2021. Moving the dial on prenatal stress mechanisms of neurodevelopmental vulnerability to mental health problems: A personalized prevention proof of concept. *Developmental Psychobiology* 63, 4 (2021), 622–640. doi:10.1002/dev.22057 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/dev.22057.

[140] Dakuo Wang, Liuping Wang, Zhan Zhang, Ding Wang, Haiyi Zhu, Yvonne Gao, Xiangmin Fan, and Feng Tian. 2021. "Brilliant AI Doctor" in Rural Clinics: Challenges in AI-Powered Clinical Decision Support System Deployment. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 697, 18 pages. doi:10.1145/3411764.3445432

[141] Liyuan Wang and Lynn Carol Miller. 2020. Just-in-the-moment adaptive interventions (JITAI): a meta-analytical review. *Health Communication* 35, 12 (2020), 1531–1544.

[142] Shijun Wang and Ronald M Summers. 2012. Machine learning and radiology. *Medical image analysis* 16, 5 (2012), 933–951.

[143] Jane C Willcox, Rosie Dobson, and Robyn Whittaker. 2019. Old-fashioned technology in the era of "bling": is there a future for text messaging in health care? *Journal of medical Internet research* 21, 12 (2019), e16630.

[144] Ping Xu, Claudia González-Vallejo, and Benjamin T Vincent. 2020. Waiting in intertemporal choice tasks affects discounting and subjective time perception. *Journal of Experimental Psychology: General* 149, 12 (2020), 2289.

[145] Wei Xu, Marvin J. Dainoff, Liezhong Ge, and Zaifeng Gao. 2023. Transitioning to Human Interaction with AI Systems: New Challenges and Opportunities for HCI Professionals to Enable Human-Centered AI. *International Journal of Human–Computer Interaction* 39, 3 (Feb. 2023), 494–518. doi:10.1080/10447318.2022.2041900

[146] Xinghui Yan, Mark W Newman, Sun Young Park, Angelle Sander, Sung Won Choi, Jennifer Miner, Zhenke Wu, and Noelle Carlozzi. 2023. Identifying Design Opportunities for Adaptive mHealth Interventions That Target General Well-Being: Interview Study With Informal Care Partners. *JMIR Formative Research* 7 (2023), e47813.

[147] Min-Jeong Yang, Steven K Sutton, Laura M Hernandez, Sarah R Jones, David W Wetter, Santosh Kumar, and Christine Vinci. 2023. A Just-In-Time Adaptive intervention (JITAI) for smoking cessation: Feasibility and acceptability findings. *Addictive behaviors* 136 (2023), 107467.

[148] Qian Yang, Yuexing Hao, Kexin Quan, Stephen Yang, Yiran Zhao, Volodymyr Kuleshov, and Fei Wang. 2023. Harnessing biomedical literature to calibrate clinicians' trust in AI decision support systems. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg, Germany, 1–14.

[149] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. doi:10.1145/3313831.3376301

[150] Qian Yang, Aaron Steinfeld, and John Zimmerman. 2019. Unremarkable AI: Fitting Intelligent Decision Support into Critical, Clinical Decision-Making Processes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3290605.3300468

[151] Dong Whi Yoo, Hayoung Woo, Viet Cuong Nguyen, Michael L. Birnbaum, Kaylee Payne Kruzan, Jennifer G Kim, Gregory D. Abowd, and Munmun De Choudhury. 2024. Patient Perspectives on AI-Driven Predictions of Schizophrenia Relapses: Understanding Concerns and Opportunities for Self-Care and Treatment. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–20. doi:10.1145/3613904.3642369

[152] Lida Zhang, Zachary King, Begum Egilmez, Jonathan Reeder, Roozbeh Ghaffari, John Rogers, Kristen Rosen, Michael Bass, Judith Moskowitz, Darius Tandon, Lauren Wakschlag, and Nabil Alshurafa. 2018. Measuring fine-grained heart-rate using a flexible wearable sensor in the presence of noise. In *2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. 160–164. doi:10.1109/BSN.2018.8329683

[153] Shanchun Zhang, Zheyuan Ding, Hui Liu, Zexin Chen, Jinhua Wu, Youding Zhang, and Yunxian Yu. 2013. Association between mental stress and gestational hypertension/preeclampsia: a meta-analysis. *Obstetrical & gynecological survey* 68, 12 (2013), 825–834.

[154] Yiran Zhao and Tanzeem Choudhury. 2024. Evaluate Closed-Loop, Mindless Intervention in-the-Wild: A Micro-Randomized Trial on Offset Heart Rate Biofeedback. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 307–312.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009