Functional Design Requirements to Facilitate Menstrual Health Data Exploration

Georgianna Lin
University of Toronto
Canada
blue.lin@mail.utoronto.ca

Pierre-William Lessard
University of Toronto
Canada
p.lessard@mail.utoronto.ca

Minh Ngoc Le
University of Toronto
Canada
ngocminh.le@mail.utoronto.ca

Brenna Li
University of Toronto
Canada
brli@cs.toronto.edu

Fanny Chevalier
University of Toronto
Canada
fanny@cs.toronto.edu

Khai N. Truong
University of Toronto
Canada
khai@cs.toronto.edu

Alex Mariakakis
University of Toronto
Canada
mariakakis@cs.toronto.edu

ABSTRACT

Menstrual trackers currently lack the affordances required to help individuals achieve their goals beyond menstrual event predictions and symptom logging. Taking an initial step towards this aspiration, we propose, validate, and refine five functional design requirements for future interface designs that facilitate menstrual data exploration. We interviewed 30 individuals who menstruate and collected their feedback on the practical application of these requirements. To elicit ideas and impressions, we designed two proof-of-concept interfaces to use as design probes with similar core functionalities but different presentations of phase timing predictions and signal arrangement. Our analysis revealed participants' feedback regarding the presentation of predictions for menstrual-related events, the visualization of future signal patterns, personalization abilities for viewing signals relevant to their menstrual experience, the availability of resources to understand the underlying biological connections between signals, and the ability to compare multiple cycles side-by-side with context.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; • **Applied computing** → **Health informatics**.

KEYWORDS

menstrual tracking, gendered health, health interface, health informatics

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642282

ACM Reference Format:
Georgianna Lin, Pierre-William Lessard, Minh Ngoc Le, Brenna Li, Fanny Chevalier, Khai N. Truong, and Alex Mariakakis. 2024. Functional Design Requirements to Facilitate Menstrual Health Data Exploration. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3613904.3642282

1 INTRODUCTION

Menstrual trackers are commonly viewed as tools that allow individuals to log their symptoms and predict when their next menstrual cycle will occur, but individuals often seek far more information about their menstrual health [14, 18, 34, 54]. They want to know and reflect on how their daily routine impacts their menstrual health and its resulting symptoms, how their menstrual health impacts their daily routine, and whether their menstrual health can be considered healthy [14, 35, 43, 54]. Answering these questions could ultimately lead to a heightened sense of bodily self-awareness that can empower individuals to make informed decisions about their overall health and wellbeing [35].

However, individuals who menstruate frequently report that they are unsatisfied with their current menstrual trackers because the limited functionality of their trackers prevents them from answering all of their questions [14]. These tools primarily serve as diaries, allowing users to revisit the data that they collected on any given day, yet they do not enable users to analyze their historical data in aggregate to identify trends and patterns over time [35]. Furthermore, existing tracking apps are primarily limited to monitoring flow timing and duration [58], which limits the lenses through which people can inspect and reflect on their menstrual health. There has also been a growing body of literature espousing the importance of accounting for the diverse and personal relationships between the menstrual cycle and physiological data (e.g., body temperature and heart rate) to comprehend individual differences in menstrual cycle characteristics [6, 12, 37, 51]. Menstrual tracker users in previous HCI studies have also hypothesized that a variety of signals from different data collection methods (e.g., physiological

data from smartwatches) may be relevant to understanding their menstrual experience [14]. Yet, less attention has been devoted to understanding how people would navigate such data and reconcile it with self-reported characteristics of their menstrual cycle.

The purpose of this paper is to provide insight into how future interface designs can effectively support users' goals in exploring their menstrual data. Previous work provides a thorough commentary on the goals that users aim to achieve, but it has lacked exploration of practical requirements to assist users in accomplishing these objectives. Since most menstrual trackers are designed specifically to provide predictions related to menstrual cycle events, users may not be able to rely solely on their previous experiences with their trackers to envision the ideal features necessary to satisfy all of their menstrual tracking goals. Instead, we use design probes to inspire users to think broadly about how menstrual trackers can better suit their needs.

To start this research, we first examined prior literature and conducted a formative study to confirm the goals users wish to achieve during menstrual health tracking. We then propose five functional design requirements aimed at facilitating multivariate menstrual data comprehension towards one or many of these goals. To validate and refine these requirements and how they could be embodied in a menstrual tracker, we gathered feedback from 30 people who menstruate. Most of these participants were young adults who, while constituting a significant user base for menstrual tracking apps [7], may not fully represent all menstrual tracking experiences (e.g., pregnancy or menopause). We then designed two mock interfaces and used them as design probes to elicit participants' ideas and impressions. During the study, participants shared how they would achieve their menstrual tracking goals with each interface separately. They then compared and contrasted the interfaces, providing suggestions for improvements in a real-time co-design activity. From our thematic analysis of study transcripts and artifacts, we provide a refined list of design requirements for future menstrual tracker designs that can effectively assist users in exploring their menstrual data and achieving their goals.

2 RELATED WORK

In our examination of prior literature, we first delineate the specific goals related to menstrual health tracking that have been identified in previous research. We then highlight the existing gaps in current platforms for effectively supporting users' understanding of their menstrual health data. Lastly, we explore how previous studies have facilitated the exploration of personal health informatics data, drawing inspiration from their approaches.

2.1 Motivations to Explore Menstrual Health Data

Personal health informatics involves the use of technology to track and analyze one's own health data to invoke positive changes in behavior, ultimately leading to improvements in mental and physical wellbeing [15, 36, 39, 40, 48]. Menstrual health informatics lies at the intersection of personal health informatics and gendered HCI, covering topics such as reproductive health [8, 9], intimate care [1, 2, 13, 57], and menopause [3, 25, 33, 55].

Various goals for tracking menstrual health data have emerged from literature in menstrual health informatics. We compiled these works and aggregated the motivations and needs of people who use menstrual trackers as described in study analyses, interview quotes, and workshop artifacts. We consolidate these goals and their sources as follows:

- **G1. Planning for future cycles:** People want to understand the timing of their upcoming menstrual phases[1] and anticipate symptom onset [14, 20, 34, 35, 54]. This is because they want to prepare for changes that coincide with transitions from one phase to the next.
- **G2. Exploring symptom patterns:** People also want to investigate their symptom patterns [14, 34, 53, 54]. They seek to identify which of their menstrual symptoms (e.g., cramps, mood swings) are most prominent over time [10, 20, 34, 35], even going as far as searching for correlations between symptoms [20, 34].
- **G3. Evaluating health status:** People strive to understand whether their menstrual cycle is "healthy" or "normal" [16, 35, 54]. This includes identifying unique characteristics of their menstrual experience from potential areas of concern that may require further consultation with medical professionals. It is important to note that this goal can be problematic, as scholars have argued that there is no simple definition for normality between and within individuals [18, 29, 31, 52].
- **G4. Expanding menstrual education:** People want to expand their menstrual literacy by learning more about how different signals are related to their menstrual health and which signals are most relevant [14, 34, 35, 54].

Although these motivations provide a valuable starting point for understanding individuals' goals with menstrual data, they fall short of providing actionable design requirements that can effectively assist users in achieving these objectives.

2.2 Supporting Menstrual Health Data Exploration

Prior research has reported that most menstrual trackers lack the functionality to fully support the aforementioned goals. For example, multiple studies have highlighted that prediction algorithms for cycle phase timing are often inaccurate and unintuitive [14, 18, 19, 58]. Lin et al. [35] found that their participants were frustrated by their inability to inspect how predictions were made, view historical data, and identify abnormalities within their logged data. Meanwhile, Epstein et al. [14] found that while some of their participants were actively tracking physiological data that may be relevant to their menstrual health (e.g., heart rate [23], body temperature [37]), they were unable to connect such data to what was being logged in their menstrual health due to the lack of cross-application integration [14]. According to Costa Figueiredo et al. [10], existing visualizations, such as calendar views, restrict users from conducting a thorough analysis of their complete cycle when their cycles lasted more than a month. In addition, graph representations solely focused on one signal which hindered users' comprehension of how

[1]For the purposes of our work, we define the menstrual cycle according to four phases: starting with menstruation, then follicular phase, ovulation, and luteal phase.

it correlated with other data they cared about. Although participants in a study conducted by Tuli et al. [54] felt more connected with their bodies while engaging with existing visualizations of their menstrual health data, many chose not to use menstrual trackers due to the absence of tools to assist in identifying menstrual health status and enhancing menstrual health literacy.

Despite the growing demand for interactive interfaces that facilitate comprehensive exploration of menstrual health information, there has been limited research on effective methods that support users' various tracking goals. Examples of efforts to both develop and evaluate new menstrual tracking solutions include work by Flemings et al. [17] and Homewood and Vallgårda [26]. Both sets of authors designed lighting fixtures that used color to indicate different stages of the owner's menstrual cycle. The main emphasis of these works was to seamlessly incorporate predicted menstrual cycle data into users' daily lives **(G1)** rather than to address how users could effectively understand symptom patterns, track the evolution of cycles over time, and increase their menstrual health literacy **(G2, G3, G4)**. As follows, our objective is to identify and validate a set of functional design requirements that can facilitate individuals in achieving all of the goals outlined in Section 2.1 when they analyze their menstrual health data.

2.3 Supporting Personal Health Exploration

Numerous studies have showcased the effectiveness of interactive visualizations and visual analytic tools in helping users comprehend and analyze large amounts of personal health data, leading to a greater sense of control that translates to comfort and potential behavior change [15, 30, 46]. Given the limited research on facilitating interpretations of menstrual health data, we examine broader studies in health informatics to inspire how we instantiate our design requirements as features in mock interfaces to elicit feedback from menstrual tracker users.

Bidwell et al. [4] and Kim [30] emphasize that the inability to observe changes in health data over time significantly hampers users' ability to identify relationships and patterns, particularly in relation to periodic data related to seasons or menstrual cycles. Some studies have suggested that layering and aggregating multiple graphs of physical activity data over time can facilitate quick comparisons across different time frames [38, 41]. On the other hand, Sharmin et al. [49] suggested utilizing summaries of descriptive statistics at different granularities so that users can obtain both high- and low-level information about their stress data. To illustrate past and predicted health-related events alongside such data, researchers have compared timeline designs that are calendar-based, linear, spiral, and circular [11, 27, 49]. Motivated by these techniques, we incorporate time-series data in our design to enable the exploration and comparison of multivariate data. We also explore different timeline designs across our design probes in order to elicit best practices in displaying menstrual cycle event predictions.

Another relevant category of health exploration work deals with leveraging data representations to encourage introspection and improve data intuitiveness. In our study, we borrow two approaches used in the literature: anthropomorphic cues and graph-based methods. Anthropomorphic cues can elicit heightened self-awareness that motivates users to strive towards desired standards [28, 47]. In the clinical space, mapping patients' health data to a human figure has improved patients' and clinicians' perception of vital sign information by reducing perceived workload and enhancing patient understanding of their health status [21, 22, 42, 45, 56]. In graph-based methods, data is structured around a root node typically representing the user or a process to communicate the attributes and relationships of personal data [44, 49]. Bosanska et al. [5] highlights the capability of such graphs to facilitate ease of information retrieval and flexibility in data representation. Drawing inspiration from the above studies, we compare and contrast design probes that organize health signals around anthropomorphic cues versus a graph-based approach.

3 METHODS

Our study aimed to gain a deeper understanding of how menstrual tracker designs can effectively support users in accomplishing their goals through the exploration of data. In this section, we outline our participant recruitment, user study, and analysis procedures. The study was conducted in three stages: (1) a brief enrollment survey to validate prior menstrual tracking goals and inform a list of functional design requirements, (2) the creation of design probes informed by participants' experiences as described in the enrollment survey, and (3) semi-structured interviews during which participants were asked to provide feedback on the design probes to validate our design requirements. The Research Ethics Board at the University of Toronto approved this study protocol under Protocol 44671. Participant responses are labeled using the notation PX throughout the paper.

3.1 Participants

We recruited 30 participants across Greater Toronto Area through social network email chains, social media groups, and workspaces affiliated with women's health advocacy organizations. Participants had to be at least 18 years old and actively tracking their menstrual cycle. The key demographics of these participants are shown in Table 1. All participants took part in both the enrollment survey and the interviews.

3.2 Stage 1: Enrollment Survey

Given the rapid evolution of health informatics tools, we aimed to confirm and expand upon menstrual tracking goals outlined by previous literature. We invited participants to reflect on their current menstrual tracking practices and goals in a survey that asked them about how long they have engaged in menstrual tracking, the type of data they tracked related to their menstrual cycle, the goals they aimed to achieve by tracking, and the challenges they faced while doing so. The survey contained 42 questions and took 30 minutes to complete on average. Findings from the survey helped us identify if our participants had any goals that were not reflected in previous work and provided explicit examples of how they wanted to achieve their goals through menstrual tracking. These insights informed a list of functional design requirements (Section 4.2).

Table 1: The demographics of our study population (N = 30).

Age	Min = 18, Max = 25, Mean = 21.2, Median = 20
Ethnicity	East Asian (11), South Asian (8), Caucasian (6), African (2), Latina (2), Middle Eastern (1)
Self-Identified Gender	Woman (30)
Education	High school graduate, equivalent, or less (1), Some university / post-secondary, no degree (18), Bachelor's (6), Master's (2), Doctorate or professional degree (2)
Years of Menstrual Tracking	Less than 1 year (1), 1–2 years (6), 3–5 years (19), 6–10 years (2)
Current Menstrual Tracker	Flo (10), Clue (8), Apple Health (5), 28 (1), Period Calendar (1), Samsung Health (1), Physical Calendar (1), Memory (1)

3.3 Stage 2: Creation of Interface Design Probes

Informed by previous work and elaborated descriptions from participants in the enrollment survey, we validated our list of menstrual-tracking goals and generated a set of functional design requirements for menstrual-tracking interfaces (Section 4). We used these requirements to guide the design of interface probes that would be shown to participants during the semi-structured interview sessions in the third stage of our research. The interfaces were built iteratively using Figma so that we could swiftly modify them during the interviews according to participant feedback. Furthermore, the interfaces were developed with a medium level of fidelity so that participants could grasp the concepts while also giving them the impression that further enhancements could be implemented. We defer the description of the probes themselves until Section 4.

3.4 Stage 3: Semi-Structured Interview Sessions

The third stage of our research aimed at validating and refining the functional design requirements. We used the interfaces we had developed as design probes during one-on-one, semi-structured interview sessions held over the Zoom videoconferencing platform and conducted by a lone researcher. The researcher displayed the design probes using Zoom's screen-sharing functionality. For each interface, participants first provided their initial impressions of the interface. They were then asked to reflect on how they would utilize the interface to fulfill each of the menstrual-tracking goals identified from our literature review and enrollment survey. Unlike in a traditional think-aloud study where participants are asked to articulate their thoughts on feature discoverability and functionality, we asked participants to specify the steps and sub-goals they want to achieve using the interface. The researcher manipulated the interface to demonstrate how it could support these steps, after which they asked participants about their reactions to this interaction, the displayed data, and their anticipated actions based on what was shown on the screen. This method of piloting the interface and interacting with participants ensured that their feedback was focused on whether the interface sufficiently supported them in achieving their menstrual-tracking goals in terms of core functionality as opposed to the details of how to practically operate a still-exploratory interface design.

We followed this process for both interfaces, alternating the presentation order between sessions. Once participants had seen and provided feedback for both interfaces, we asked them to compare and contrast the functionality and visual aspects of concepts. We then invited them to provide feedback on how they would improve one of the interfaces to better suit their needs. Participants were free to choose the interface they would provide feedback on according to preference or the volume of feedback they had to provide. This part of the session was treated as a co-design activity; participants provided suggestions to the researcher, and the researcher made design changes in real time. Each session took approximately 60 minutes, and participants were given $40 CAD for completing both the survey and the interview session. Any changes that were made to the interfaces were noted and reset between sessions.

3.5 Analysis

Three researchers processed the survey responses through open coding, and they resolved conflicts and missing codes through discussion until a consensus was reached. The codes were grouped using thematic analysis [6], resulting in a set of menstrual-tracking goals that were compared with those identified from the literature.

The same researchers analyzed the semi-structured interview sessions in a similar manner. All of the semi-structured interview sessions were recorded and transcribed, after which open coding and thematic analysis were used to examine the session transcripts, notes, and artifacts. These themes included but were not limited to "explicit calendar date predictions", "personalize relevant signals", and "textual explanation of biological connection".

3.6 Positionality

The researchers involved in this study come from diverse academic backgrounds involving gendered health, ubiquitous sensing, visualization for health, qualitative research, and HCI research broadly. The research was conducted in regions of North America that may exhibit biases aligned with WEIRD characteristics (Western, educated, industrialized, rich, and democratic) [24]. Although the authors have varied cultural backgrounds and upbringings from different parts of the world, the study's findings predominantly resonate with Western perspectives on feminism.

4 DESIGN REQUIREMENTS AND PROBES

Below, we describe how the results of our enrollment survey supported the menstrual-tracking goals we identified from previous

literature (Section 2.1). We also report on the perceived shortcomings of existing menstrual trackers that participants noted in the same survey. We use these findings to propose an initial list of functional design requirements that designers can follow to create menstrual-tracking interfaces. Finally, we describe two design probes that we created to validate and refine these requirements during our semi-structured interview sessions.

4.1 Validation and Expansion of Menstrual-Tracking Goals from Participants

Participants' survey responses aligned with the menstrual-tracking goals we compiled from prior work. Furthermore, they provided tangible and extensive insights into the motivations behind these goals. For example, most participants (N=27) stated that they rely on menstrual trackers to anticipate the timing of their next menstruation **(G1)** and the potential symptoms associated with it **(G2)**. They used this information to not only anticipate body changes but also to take preparatory action to manage symptoms. Some participants mentioned that they would buy pads to manage their menstrual flow, and others stated that they planned to eat different meals to prevent acne.

> *"I want to plan out my exercise routines, be prepared with pads or any medication when I'm out of the house, anticipate my cravings and cook more when I'm about to get my period."* (P23)

On the other hand, participants noted that current trackers often do not provide symptom predictions (N=11), which limited their ability to anticipate and correlate symptoms with their menstrual cycles in real time **(G1, G2)**. Additionally, they did not feel that current trackers were adequately equipped to facilitate the explorations they desired to undertake with their historical symptom data **(G2)**. For instance, P3 mentioned that her current menstrual tracker lacks a feature to visualize her historical information; rather, it only presented data from a single day selected from a calendar.

> *"I have to manually go through all the pairs [of days] to see if I notice anything. I used to track a lot of things, but that information never came back to me. [With my current tracker], I'm not gonna learn anything outside of like my own brain, or I guess I could go back and scrape the data and do an analysis. But that's a lot of work."* (P3)

With regard to health status **(G3)**, many participants (N=21) were particularly concerned about recognizing differences between their menstrual cycles over time since they perceived unexpected changes in their cycle characteristics to be potential signs of concern. Participants defined cycle consistency according to multiple dimensions, including both menstruation characteristics (e.g., timing, blood volume) and symptom characteristics (e.g., occurrence, severity). However, participants noted that they faced difficulties in determining cycle consistency using current applications since they were unable to examine other cycles for a frame of reference.

> *"I have to rely on memory while swiping through the months to compare with other months."* (P6)

With regard to education **(G4)**, participants (N=16) hoped to become more aware of their bodies and exchange information with their peers. This required examining not only self-reported data but also physiological signals that may or may not be associated with their periods (e.g., heart rate, exercise). Yet, participants felt that current applications lacked the capability to record and analyze many of the signals they wanted to track.

> *"I believe my app could be even better with a personalization option, leveraging additional data from my Apple Watch, such as heart rate, exercise level, sleep duration, and more."* (P26)

4.2 Functional Design Requirements

Having validated people's menstrual-tracking goals and identified ways in which menstrual trackers fail to support them, we further synthesized our findings to map out the steps that people would ideally take in order to achieve their goals. The diagram shown in Figure 1 describes the intertwined processes that participants said they would want to follow to accomplish any of the four menstrual-tracking goals that resonated with them. The five actions labeled in magenta are steps that participants are already able to complete or wish that they could complete with their menstrual trackers. Therefore, we deem these actions to be functional design requirements and elaborate upon them below:

- **DR1. Cycle Event and Symptom Predictions:** The interface should effectively communicate the predicted timing of menstrual cycle phases, symptom onset, and events like the onset of menstruation and ovulation **(G1)**.
- **DR2. Support for Multivariate Data:** The interface should support a variety of health signals, including dates of menstrual cycle events, self-reported symptoms, and physiological data **(G2, G3, G4)**.
- **DR3. Inspection of Individual Signals:** The interface should offer varying levels of detail (e.g., raw data, daily summaries, cycle summaries) for individual signals **(G2, G4)**.
- **DR4. Comparison of Multiple Signals:** The interface should show how multiple signals vary with or in contrast to one another across the menstrual cycle **(G2, G3, G4)**.
- **DR5. Comparison Across Cycles:** The interface should enable users to compare various cycles, including their own historical cycles and those of their peers **(G3, G4)**.

4.3 Interface Probes

We developed two contrasting design probes to produce concrete instantiations of the design requirements, enabling us to solicit feedback during our semi-structured interviews. Our aim with these probes was to elicit thoughts and ideas about the functionality of the proposed designs rather than the specific design decisions themselves. The data shown in both interfaces were simulated using common patterns presented in prior menstrual health literature [6, 23, 37, 50]. Participants were explicitly informed that the data was not based on any specific individual. However, they were asked to react to it as if it were their own.

The two interfaces, called MenstrualMate and PeriodBubble, are shown in Figure 2 and Figure 3 respectively. We designed MenstrualMate to evoke a personal connection between the user's data and their body, showcasing signals on a human figure and events on a linear timeline. Meanwhile, we designed PeriodBubble to be

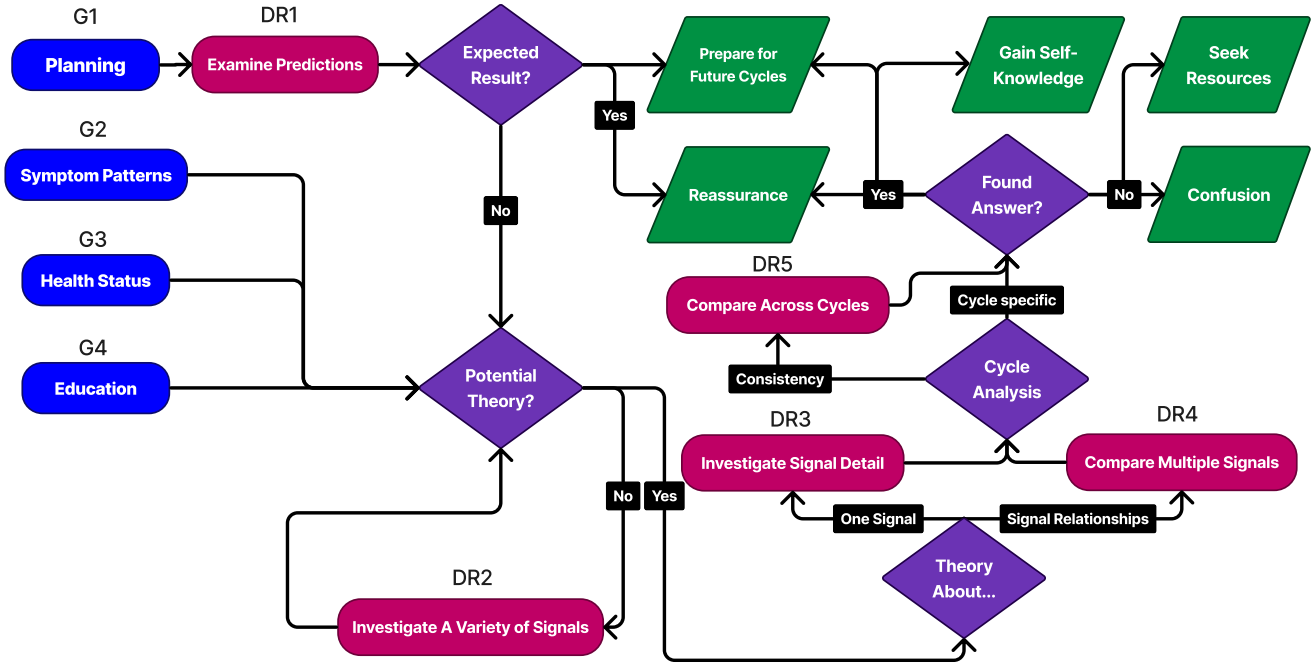

Figure 1: A flow chart describing participants' ideal decision-making process for accomplishing their menstrual-tracking goals. The goals (G1–G4) are shown in blue rounded rectangles, the actions informing our design requirements (DR1–DR5) are shown in magenta rounded rectangles, the decision points are shown in purple diamonds, and the outcomes are shown in green parallelograms.

more data-centric, presenting signals as nodes around a central hub with a circular timeline. Below, we elaborate on how both interfaces satisfy our design requirements in similar and contrasting ways.

4.3.1 Predictions (DR1). To examine the timing of past and future menstrual cycle events, both interfaces provide a timeline showing transitions across menstrual cycle phases starting from the first day of menstruation to the last day of the luteal phase. Since viewing predictions is a frequent and crucial action for menstrual tracker users, we opted for two different timeline designs to understand how participants might react similarly or differently to different temporal representations. Both timelines have the capability to accommodate variable lengths, allowing users to view a complete cycle's worth of data. In PeriodBubble, menstrual cycle timelines are presented in a circular manner within large hub nodes on the outer sides of the visualization (Figure 3, A) to convey the cyclic nature of the menstrual cycle. In MenstrualMate, menstrual cycle timelines are presented linearly along the top and bottom of the screen (Figure 2, A), incorporating calendar dates to align with planning schedules. The examples in the figures indicate that the user is on the tenth day of their current menstrual cycle. Both interfaces highlight the current day within the menstrual cycle using a pronounced circle overlaid on their respective timelines (Figure 2 and Figure 3, A). Both interfaces also provide red and blue highlighting to indicate the menstrual and ovulation phases, respectively. However, MenstrualMate exclusively provides text labels for all four menstrual cycle phases to elicit feedback on whether participants would find all phases equally important to support menstrual health reasoning.

To provide richer predictions beyond menstrual cycle events, each interface includes support for predicting future health signal measurements. Both interfaces include selectable icons representing the signals available to users (Figure 2 and Figure 3, B). Once selected, the signals can be dragged onto the central chart (Figure 2 and Figure 3, C), where the signal's data will be plotted as a solid line. The x-axis presents temporal progression in two ways in order to encourage conversation: (1) the number of days since the start of the selected cycle and (2) a normalized percentage through the cycle. Future predictions of that data are shown as a dashed line that extends beyond the current date (Figure 2 and Figure 3, C).

4.3.2 Support for Multivariate Data (DR2). When the user first sees either interface, they are instantly presented with a rich range of signals related to their menstrual health (Figure 2 and Figure 3, B). The list of signals we included in our interface probes included menstrual cycle phase timing, self-reported symptoms (acne, cramps, sleep levels, and exercise levels), and continuously recorded physiological data (resting heart rate and body temperature). This list was inspired by a combination of existing health trackers and clinical literature related to the physiology of menstrual health [6, 14, 23, 37], with a preference towards signals that are readily measured by wearables and consumer-grade devices. In PeriodBubble, these signals

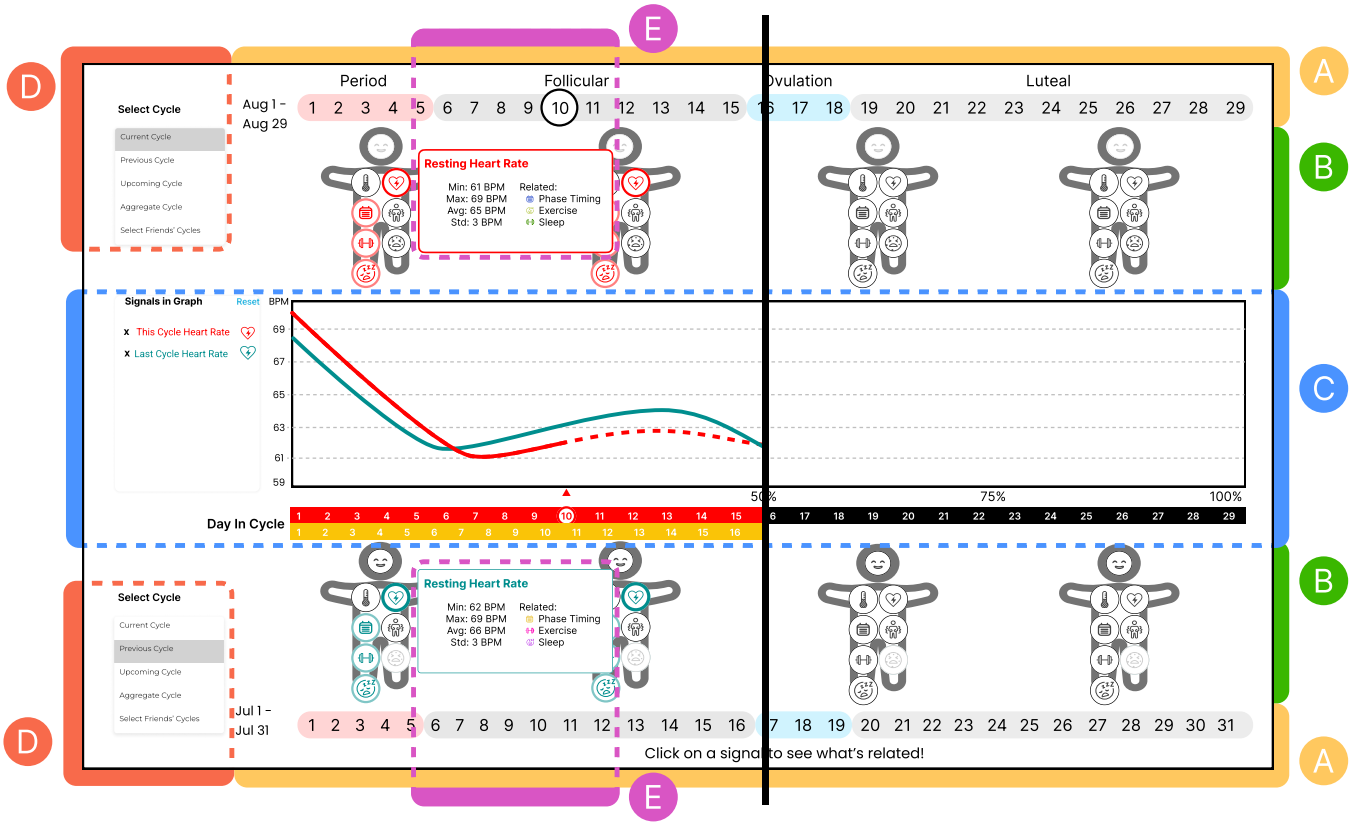

Figure 2: The right side of the figure shows MenstrualMate in its default state without any data being shown, while the left side of the figure shows the interface populated with heart rate from the user's current menstrual cycle and temperature data from their last cycle. MenstrualMate has the following notable features: (A) linear timelines showing past and predicted menstrual cycle phases, (B) human figures for each phase with selectable signal icons, (C) a central chart where users can view past and predicted data, (D) a dropdown list to select different menstrual cycles for comparison, and (E) a box showing summary statistics for the plotted signals.

are presented in smaller nodes around the central hub. In MenstrualMate, the signals are illustrated on human figures (Figure 2, B) such that the spatial arrangement closely matches the part of the body most associated with them. For example, mood is depicted on the face, and heart rate is depicted over the upper half of the chest. We chose to differentiate the two probes in this way to not only encourage conversations on signal presentations but also to assess whether placing signals on the body would enhance participants' understanding of the type of data they are collecting. Since adherence to self-tracking practices is not always continuous [32], we explicitly indicate when data is missing by leaving the corresponding icon semi-transparent; in our examples, the previous cycle's acne and the current cycle's mood data (Figure 2 and Figure 3, B)

*4.3.3 Signal Inspection and Comparison **(DR3, DR4)**.* The user can click on signal icons within both interfaces to reveal a box that presents summary statistics and historical trends related to that data (Figure 2 and Figure 3, E). The presented statistics include the signal's minimum, maximum, and average signal values across the given cycle.

In PeriodBubble, the user can only select icons corresponding to data across an entire menstrual cycle. In MenstrualMate, however, each signal icon represents a single phase of the menstrual cycle (Figure 2, B). Furthermore, the summary boxes for self-reported symptoms across both interfaces identified the menstrual cycle phase during which the user experienced the highest severity levels. These differences were intended to probe the granularities at which people required to investigate summaries and trends in their data.

When the user selects a signal, the other signal icons that have a relationship to the selected signal are shaded in a lighter color that matches the selected signal's hue. The user can directly compare and correlate signals by dragging them into the empty chart in the middle of both probes. The left side of Figure 2 shows the interface after selecting the resting heart rate signal and dragging it onto the chart. Within this example, users could click and drag other signals within either cycle onto the chart to compare and correlate their trends alongside the resting heart rate signal. Signals with similar ranges of possible values (e.g. two symptoms with 7-point Likert ranges) are given a shared y-axis, whereas signals with different ranges are given independent y-axes on either side of the chart.

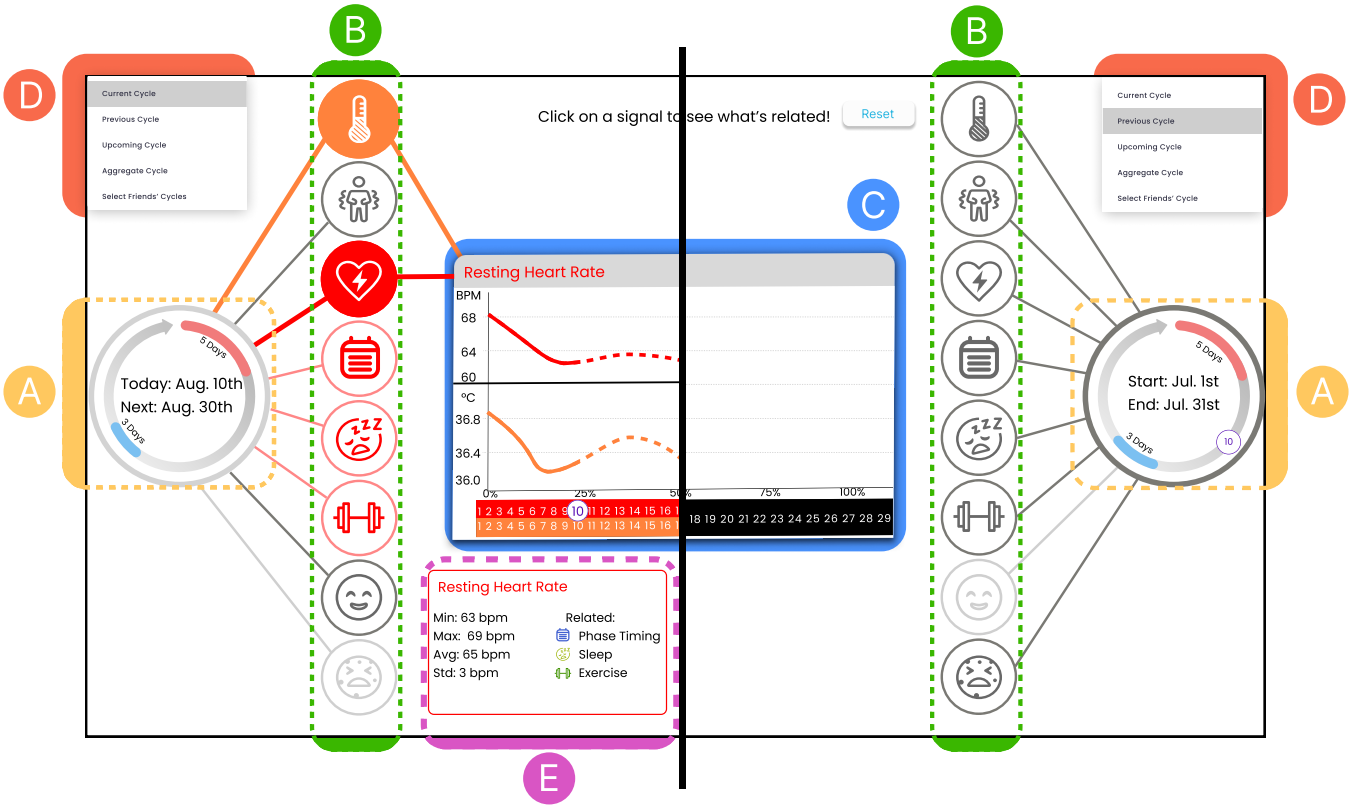

Figure 3: The right side of the figure shows PeriodBubble in its default state without any data being shown, while the left side of the figure shows the interface populated with heart rate and temperature data from the user's current menstrual cycle. PeriodBubble has the following notable features: (A) circular timelines showing past and predicted menstrual cycle phases, (B) selectable signal icons as graph nodes, (C) a central chart where users can view past and predicted data, (D) a dropdown list to select different menstrual cycles for comparison, and (E) a box showing summary statistics for the signals being plotted.

4.3.4 Comparison Across Cycles (DR5). Beyond comparing different signals within their current menstrual cycle, the user can also compare signals across aggregates of other menstrual cycles. The user can select individual and aggregated cycles using the dropdown menus at the corners of the interfaces (Figure 2 and Figure 3, D): their current cycle, their previous cycle, their upcoming predicted cycle, the aggregate of all of their past cycles, a friend's current cycle, and the aggregate of their friends' current cycles.

Just as how the user can compare signals within their current cycle, they can also compare signals across cycles by dragging icons into the chart in both probes. In fact, the user can compare up to four signals within and across cycles at the same time to see how correlations evolve over time. For example, the user may want to examine the correlation between exercise and self-reported cramps within their peers' data and their own data at the same time.

5 FINDINGS

5.1 Cycle Event and Symptom Predictions (DR1)

Participants (N=20) greatly valued seeing predictions of when they were likely to menstruate. By aligning their cycle events with their schedules, participants speculated that they would be able to better plan and manage their routines with greater confidence knowing when their typical day-to-day happenings might be impacted.

> *"If I go on a trip, …I just want to see the red part [days of menstruation] compared with my itinerary for those days."* (P16)

Participants preferred the linear timeline in MenstrualMate compared to the circular one in PeriodBubble since it was easier for them to envision how the former would align with their calendar. They also appreciated the percentage-based annotations indicating the progression through the menstrual cycle for similar reasons. When they saw that the percentage reached a high value, they knew that they would be approaching a transition between cycles and, therefore, a transition to a new menstrual phase. However, some participants noted that they would want to see progression expressed according to other time frames. For example, P30 suggested that percentage-based annotations should be framed according to phases (e.g., "50% through the luteal phase") rather than entire menstrual cycles since she believed that phases could be more precisely defined, unlike cycles that do not have definite start and end dates.

Following the notion that menstrual cycles do not have fixed endpoints, participants commented that the circular timeline used

in PeriodBubble aligned better with their perception of menstrual cycles. A few participants (N=4) explicitly stated that while linear representation of cycles might be easier to look ahead in their schedules for planning, the circular design resonated more with their personal understanding of menstrual cycles and made it easier for them to reflect on their bodies. For these reasons, P13 said that they imagined using MenstrualMate on a daily basis or as frequently as needed to utilize their menstrual data for planning purposes, but she would also examine PeriodBubble once every few cycles to reflect on past data. P3 had the following to say about how she would use the circular timeline for reflection:

> *"I would look once a month, or once every couple of months to see the data in aggregate . . . If I notice I'm really moody towards the beginning of my cycle or the end of my cycle, I'd make more of an effort to think that it might not be because of the situation."* (P3)

Participants also valued the representations employed in the interface to convey the uncertainty of predicted symptom trends. The use of dashed lines and measures of cycle progression rather than calendar dates in the central chart was found effective at communicating the hypothetical nature of the symptom predictions while enabling some level of planning. Participants believed that future symptoms change more frequently compared to menstrual cycle events. Consequently, including dates for every change would result in unnecessary clutter and make it challenging to comprehend how the symptoms might vary across their projected cycle. Participants (N=27) envisioned regularly checking how their symptom predictions fluctuated in the chart to optimize daily tasks and prepare for anticipated discomfort such as cramps and headaches.

5.2 Support for Multivariate Data (DR2)

Participants were pleasantly surprised and excited upon discovering the diverse range of signals included in both interfaces, as their current menstrual trackers typically only supported self-reported symptoms and did not incorporate any physiological data collected by other devices.

> *"I do track my sleep on my phone, but it's not connected to my cycle app. It's just two separate trackers . . . I can't relate them. So connecting those two is very useful."* (P8)

Although participants considered most of the signals familiar and relevant, they felt that some signals specific to their own menstrual experiences were missing. Therefore, they expressed interest in a feature that would allow them to add even more signals and bring them to the forefront of the interface.

> *"I think some people might have reasons for not wanting to put in like food or water intake, and I think that's totally fine. But I do think it'd be good to have it as an optional feature there because I think, for me and for lots of people, it could matter."* (P5)

Initially, participants thought they would be only interested in exploring signals they felt were related to their menstrual experiences. However, discussing how they would interact with the interfaces gradually piqued their interest in uncovering the potential relevance of signals that were less familiar or initially perceived as less significant to their menstrual health. Participants eventually saw the inclusion of unfamiliar signals as additional opportunities for learning even if they were not shown at the forefront of the interfaces.

> *"I feel like to some degree all these things are interconnected . . . I know you've got to choose the ones that make the most sense for what purpose [I have now] but also support any other purpose I would want to [examine later]."* (P5)

Few participants (N=4) mentioned that they might not explore these signals often, but they would consider setting aside time in the future to determine if these signals held any relationship to their menstrual health and if it would be worthwhile to track them more frequently. They hypothesized that putting in this effort to assess the importance of different signals to their own health and menstrual-tracking goals could streamline future data collection and investigation. As P23 shared, *"If I record my temperature and it's related to menstrual health, I might as well record [other] related signals and see what they have to do with each other."* (P23).

Participants opined on how signals were presented, but their preferences were largely centered around their ability to effectively organize and prioritize them. They noted that the added space around the nodes in PeriodBubble could grant them the ability to rearrange the nodes according to their perceived relevance. For example, some participants suggested that they would want to group signals into categories and isolate the signals most relevant to their questions. Potential groupings included self-reported symptoms versus physiological data and signals they believed to have control over (e.g., exercise) versus signals they perceived to be more challenging to disrupt (e.g., body temperature). Participants also proposed the option to hide certain signals that they deemed less relevant as a way to reduce clutter on the interface.

> *"Let the user choose the order of the little icons on the [Menstrual Mate] interface. Right now, the ones that I care about most are placed last. I would want to see cramps first."* (P16)

5.3 Inspection of Individual Signals (DR3)

Participants were told that our design probes presented data corresponding to each day of a given menstrual cycle, but they sought multiple levels of detail when examining their data. They appreciated the default of seeing data aligned with the days of a cycle because it allowed them to determine if events from a single day had a long-lasting impact on their menstrual cycle. However, they were also interested in viewing data that was smoothed and summarized at different scales because they felt that identifying trends within phases and entire cycles would allow them to answer different questions related to their health status. Participants wanted to know if data changed alongside menstrual cycle phase transitions and if there were common data patterns across menstrual cycles. For instance, P5 wanted to know how her daily sleep quality varied and whether this variance differed across menstrual phases. To accomplish this goal, she anticipated examining how her sleep quality changed across days within the same cycle If she identified any noticeable changes, she imagined examining the average of the daily fluctuations grouped by menstrual cycle phase to ascertain if the changes were more prominent during specific phases.

The importance of being able to identify signal fluctuations across menstrual cycle phases was particularly salient, as participants believed it would help them determine how signals change in conjunction with the hormonal processes of their menstrual cycle.

> *"[I want to know] what are my hormones, and how they fluctuate throughout my cycle. Then based on their fluctuation, [I want to know] how it affects other factors that I'm tracking."* (P10)

Many participants (N=16) preferred how this affordance was supported in MenstrualMate. They found the human figures in that interface to be more personable, serving as convenient markers for reflecting on different stages of their own menstrual experiences. They resonated with the idea of seeing the figures dynamically transition alongside them throughout their menstrual cycles, fostering a sense of connectedness and relatability. In turn, they believed it would be easier for them to reason how signal patterns might change between people or the phases to which they are anchored.

> *"It humanizes the experience. Each of the people represents you going through each of the different stages. A person's going through them. . . . It's personal . . . I feel a connection."* (P5)

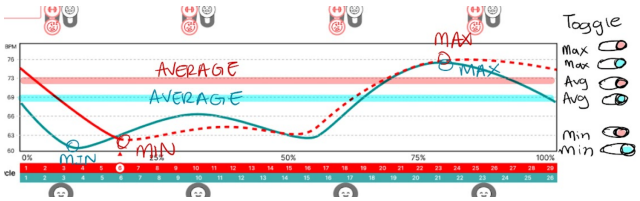

Figure 4: A drawing by P2 during the co-design phase of their interview, during which they suggested the ability to toggle annotations of important summary statistics on the central chart.

Beyond seeing data displayed in the central chart, participants found value in the presentation of summary statistics — averages, standard deviations, maximums, and minimums — to understand changes in a specific signal across the menstrual cycle. These numbers enabled them to quickly identify ranges of typical values for a particular signal that would ground the interpretation of future observations. A benefit participants (N=15) voiced about MenstrualMate was the potential for multiple text boxes to include summary statistics corresponding to each menstrual cycle phase figure. They suggested that having detailed summary statistics for each phase would allow them to easily establish differences between phases. P2 and P8 proposed that summary statistics be presented as toggleable annotations directly on the central chart (Figure 4) or as pop-ups that appear whenever a line is selected within the chart.

5.4 Comparison of Multiple Signals (DR4)

Participants liked being able to see how signals interacted with one another across the menstrual cycle. They sought to not only identify signals that were directly correlated with one another but also instances where one signal appeared to influence another with a time lag. In finding these relationships, participants hoped to surface causal relationships that would lead to actionable changes in how they manage their symptoms, develop healthier habits, address medical concerns, or seek reassurance.

> *"It would give my life a lot more meaning if I knew what happened in my body. For example, do different body temperatures during different phases in my period affect my workout? Or what are some things that I ate during certain period phases during this past year, and how did my body react?"* (P8)

> *"Mood obviously can change, sleep can change, but I want to know [if there is a] sweet spot between how much exercise I should be getting and how much sleep I should be getting to make me feel the best during my cycle."* (P7)

However, participants recognized that they could not always determine causality by simply visualizing signals on the chart and seeing potential correlations. P3 noted that two signals appearing to be statistically correlated may not have a biological link between them since the correlation may be due to random chance or limitations in how the data was recorded. She stated that knowing two signals were statistically correlated was not enough; rather, she wanted to understand the underlying mechanisms that make them related if there was a real connection between them.

P3 and many others (N=11) expressed the desire for written explanations for all potential relationships to help them determine if the corresponding signals are relevant to their health. P7, P8, P10, P12, and P23 suggested pop-ups with text explaining or refuting potential biological connections between signals. Example explanations included *"Cramps are related to the amount of exercise that you do throughout your cycle. Because you did this amount of exercise this cycle, you will experience less cramps."* (P8) or *"This signal is not related to your menstrual cycle. It's related to your diet like eating too much junk food."* (P6).

P1, P3, P7, and P10 proposed adding a hyperlink with each signal that would direct users to a separate page containing more information about them. Suggestions for resources included clinical guidance on the signal's average relationship to the menstrual cycle phases and other signals. Participants (N=7) were also interested in learning about terminology that is commonly associated with each signal, such as the differences between "resting heart rate" and "heart rate" and "basal body temperature" and "body temperature".

5.5 Comparison Across Cycles (DR5)

Participants highlighted the importance of comparing signals across multiple menstrual cycles to understand potential causes of concern, track long-term changes in their bodies, and explore historical data. Therefore, they were excited by the ability to view, compare, and contrast data across multiple cycles simultaneously. P3 noted that she would compare timelines for different cycles to understand if their duration differed, and P26 envisioned dragging and dropping the same signal icon from different cycles into the central chart to determine if they exhibited similar characteristics during a given menstrual cycle phase.

Nevertheless, some participants suggested different ways of viewing more complex relationships. P7 gave the example of investigating if the relationship between body temperature and heart rate

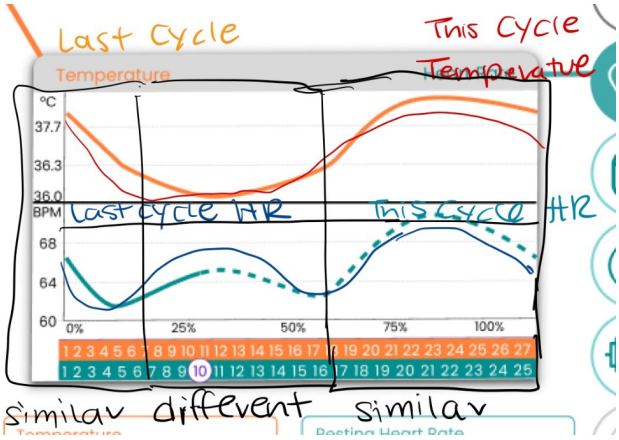

Figure 5: A drawing by P7 during the co-design phase of their interview, during which they envisioned how they would want to see two pairs of signals visualized in charts that facilitate comparisons.

persisted between cycles. They wanted to be able to drag the four corresponding icons onto the central chart in order to generate two stacked charts as shown in Figure 5: one for temperature and one for heart rate. P7 presumed that this configuration would have made it easier for them to determine how the signals compared across the different cycles.

> *"I would use the phase timing to compare the previous cycle to this cycle and see . . . if I'm getting relatively the same amount of sleep and exercise. And then, if I'm getting more painful cramps this cycle compare it to the last cycle, I can figure out what I'm doing differently between two cycles . . . like if my daily life has changed."* (P1)

We originally prompted participants with the idea that they could compare their current menstrual cycle with either recent cycles, current cycle from peers, or aggregate cycles over all of their historical data or all of their peers for any purpose. However, participants gravitated towards comparisons between menstrual cycles in the past year and their most recent cycle to assess their current menstrual health status.

> *"If I found that the most recent cycle is very different from the past few, . . . I would reflect on my past four weeks and see what made this difference or if anything impacted this difference."* (P26)

On the other hand, participants hypothesized that they would turn to data from over a year ago to gain awareness of how their menstrual experiences may have evolved as they entered different life stages or changed lifestyles.

> *"My flow is a lot lighter than like 2 years ago. Is this normal, or did it just become less heavy? I feel like, maybe after you have a kid, your period will probably change in some way, so having different years of data could kind of help."* (P6)

If these transitions were perceived as related to a certain event (e.g., moving, exams, entering college), participants suggested adding the ability to mark these events in the interface. Seeing significant events marked on the signal charts would enable them to compare data before and after to investigate the impact of these events on the stability of their signals and their menstrual health more broadly.

To help people investigate hypotheses related to their historical data, P3 and P4 proposed giving users the flexibility to select temporal ranges of cycles to be aggregated for comparison. P19 and P22 took this idea further and proposed the ability to group past cycles according to different dimensions. With this feature, they anticipated identifying common signal trends within specific cycle categories and detecting deviations in individual cycles. For instance, by separating shorter menstrual cycles from longer ones, users would be able to search for possible factors that may influence the duration of their menstrual cycle. While we normalized the temporal axis of signals across different cycles such that time was expressed as a percentage through the menstrual cycle, P3 proposed that it would also be helpful to give people the option to normalize the length of each phase so that they could more readily identify potential causes for phase length differences.

> *"If my period came a whole week early, I'd be really interested in, like what else changed? Was the whole cycle shortened by like 25%? Or was some phase cut short? Does it look like everything was the same until like this point, and then it suddenly accelerated."* (P3)

Most participants (N=16) were primarily interested in understanding changes within their own historical data since they believed that other people's menstrual data would not accurately reflect their own experiences. Nevertheless, few participants (N=9) envisioned select situations when they would find value in comparing their own aggregated data to that of their peers. For example, they expressed the desire to view the signals that had the highest impact on their peers' cycles. Under the assumption that they were already aware of their peers' lifestyles and context, participants felt that such investigations would allow them to assess the relevance of those signals to their own menstrual health. Still, P2 and P21 felt that detailed information on their peers' cycles would yield both too much data to explore and a violation of their peers' privacy, so they only wanted to see summaries of aggregated per-cycle data.

Beyond inspecting data from their peers, participants (N=10) also desired the ability to compare their trends and correlations to those of broader populations. Participants expressed curiosity about how similar or different their own cycles were compared to other people using menstrual trackers.

> *"I'd want to see a general trend. Not like, 'What is Sally on a Tuesday doing for a period?', but instead, 'This many people experience this during their luteal phase'."* (P2)

> *"I don't really see period apps often compare my symptoms to other women's symptoms. For example, when I'm cramping, it would be nice to know what percentage of women experience that and how common my symptoms are, because I feel like that is also a good way to know if you're okay."* (P17)

P8 even proposed a feature to compare their data to segments of the population filtered according to lifestyle and demographics to enable comparisons with groups who had similar characteristics. Suggestions for grouping variables included age, ethnicity, height, weight, and city. P26 elaborated on this thought further:

> *"I also want to know if I am similar to others. But, for example, I think it's not a good idea to compare my cycle with a person who is like 20 years older than me."* (P26)

If participants were to identify significant differences from others, they suggested that they would want to see pop-ups or hyperlinks to external resources explaining the range and patterns of healthy population-level baselines.

6 DISCUSSION

Results from our study suggest that the functional design requirements we articulated could help participants achieve their goals for menstrual tracking. However, participants also expanded on these requirements by providing additional detail on how they would want these requirements realized. Below, we summarize our key findings and reflect on the lessons learned from them to refine the initial design requirements. Furthermore, we justify our methodological decisions and discuss the promising opportunities for further investigation.

6.1 Cycle Event and Symptom Predictions (DR1)

Participants valued seeing menstrual cycle event and symptom predictions since it would allow them to account for their menstrual cycle in their daily schedules **(G1)**. Viewing predictions along a linear timeline enabled them to grasp how their schedule may be impacted by impending menstrual events. However, they found that circular timelines captured the recurrent nature of menstrual cycles, making it easier for them to reflect on periodic trends **(G2)**.

This highlights a tension between linear anticipation of future events and cyclical reflection on recurrent patterns. Underlying this tension is a difference in the cadence at which participants anticipated performing these investigations. They suggested that they would want to anticipate future events briefly but multiple times within the same menstrual cycle. In contrast, participants forecasted that they would reflect on recurrent patterns during a longer session once every few cycles. Since both processes were equally important to participants, designers should consider complementary or unified solutions that satisfy both framings of progression through the menstrual cycle.

Additionally, participants felt that timelines alone were insufficient for viewing future symptom predictions that were hypothetically reported along ordinal scales in our probes **(G2)**. Knowing predicted dates of when symptoms might occur was not enough; they also wanted to know how symptom severity would fluctuate over time. Participants suggested that symptom predictions should be displayed with historical context and with more uncertainty in order to foster transparency and prepare users for varied outcomes. Considering these suggestions, designers should explore strategies to support both date-driven planning (e.g., planning around menstrual events, reconciling menstrual events with other activities) and trend forecasting (e.g., projecting and managing symptoms).

- **Refined Design Requirement:** The interface should effectively communicate the predicted timing of menstrual cycle phases, symptom onset, *symptom variance,* and events. *In doing so, the interface should address varying interaction patterns and the potential tension users may experience between their linear reasoning for planning ahead and their cyclical understanding of the menstrual cycle. Furthermore, the interface should accommodate both date-driven planning and trend forecasting purposes* **(G1, G2)**.

6.2 Support for Multivariate Data (DR2)

Since most menstrual trackers do not include the rich physiological signals that we showed in our probes, participants felt that their current menstrual trackers have been preventing them from exploring potential causes of their symptoms **(G2)**. Participants also felt that incorporating multivariate data would allow them to expand their understanding of the signals they consider pertinent to their menstrual experience **(G4)**. They speculated that by identifying the significance of unfamiliar signals and signals that they did not expect to be related to their menstrual cycle, they could prioritize and streamline their future data collection and investigation efforts.

To assist in navigating diverse signals and symptoms, participants desired the ability to group and prioritize them based on their perceived menstrual health relevance. We provided this functionality to some extent, highlighting statistically correlated signals as they were moved towards the central charts. However, participants ascribed relevance along other dimensions Some participants assigned relevance according to their curiosity and intrigue, with unfamiliar signals like glucose level and body temperature being deemed more intriguing. Other participants assigned relevance according to the signals' actionability; in this regard, signals associated with behaviors they believed could be controlled were viewed as more actionable. Participants believed that sorting signals according to their perceived actionability would enhance their ability to prioritize investigations related to menstrual planning, symptom management, and habits that may influence their menstrual health **(G1, G2, G3)**. Although signals like sleep duration and activity level can be acted upon to yield health benefits, some researchers have criticized the idea of "treating" or "fixing" menstrual health [25, 54]. Therefore, designers should avoid framing actionability in this light.

Beyond giving users the ability to group and sort signals and symptoms, designers could include features that periodically surface unfamiliar menstrual signals to the forefront so that users build a comprehensive understanding of their menstrual health. Designers could also prompt users to reflect on the degree to which their decisions may or may not impact these signals, helping them avoid unrealistic notions of "fixing" aspects of their health that may not require intervention.

- **Refined Design Requirement:** The interface should support a variety of health signals, including dates of menstrual cycle events, self-reported symptoms, and physiological data. *Users should be able to personalize the grouping and arrangement of these signals according to their preferences, including factors like perceived relevance and actionability. However,*

designers must also exercise caution when conveying actionability to avoid misleading users, as many self-reported aspects of health are beyond individuals' control. **(G2, G4)**.

- **Optional Design Refinement:** *Users should also be encouraged to examine unfamiliar signals and reflect upon the extent to which they could influence them to potentially expand their menstrual health understanding* ***(G4)***.

6.3 Inspection of Individual Signals (DR3)

Although not all participants were familiar with the different phases of the menstrual cycle, they recognized that their cycles were not solely defined by moments of bleeding. Participants wanted to examine fluctuations of individual signals throughout the menstrual cycle in order to quantify the lasting impact of external events and circumstances and determine whether they were influenced by menstrual hormonal processes **(G2, G4)**. Echoing prior work [35], participants were not concerned about bodily changes that they believed to originate from hormonal processes since they considered those changes to be outside their control. Instead, they were more concerned about changes they could not attribute to their menstrual cycle, as these were believed to be linked to external factors they could influence. Hence, users may appreciate menstrual trackers that help them tease apart which symptoms might have stronger or weaker connections with menstrual hormones.

- **Refined Design Requirement:** The interface should offer varying levels of detail (e.g., raw data, daily summaries, phase summaries, cycle summaries) for individual signals. *The interface should also help users tease apart which signals may be closely tied to their menstrual hormones versus other physiological and behavioral factors* **(G2, *G3*, G4)**.

6.4 Comparison of Multiple Signals (DR4)

Participants sought to not only identify direct correlations between signals but also instances when one signal appeared to influence another **(G2)**. However, they acknowledged that they may lack the extensive medical knowledge necessary to determine the validity or causality of the relationships they identified **(G4)**. Therefore, they wanted to see textual descriptions describing biological mechanisms that may explain potential correlations. As with the inspection of individual signals, these biological mechanisms could involve hormonal processes of the menstrual cycle to help users identify relationships that are closely linked with their menstrual health. By enabling users to explore their menstrual data at different time scales with accompanying insights from literature, they may express heightened motivation to engage in exploratory analyses that unveil multivariate trends. If prior work has not yet yielded a biological explanation for a certain relationship, designers should acknowledge that current literature may not entirely account for the presence or absence of a relationship. Additionally, they could emphasize that lacking clinical commentary on the signal relationship should not discredit their observations or experiences.

- **Refined Design Requirement:** The interface should show how multiple signals vary with or in contrast to one another across the menstrual cycle. *It should point users to resources on biological mechanisms that may explain potentially causal relationships between signals. In cases where these causal relationships or biological mechanisms lack clinical literature support, the interface should convey uncertainty without confirming or dismissing their existence* ***(G1, G2, G3, G4)***.

6.5 Comparison Across Cycles (DR5)

For a comprehensive understanding of their menstrual health status **(G3)**, participants appreciated the ability to contextualize their current menstrual data relative to different groupings of their past cycles. Participants sought to establish whether their menstrual cycles adhered to a specific pattern when they were feeling healthy. Participants believed viewing cycles with inconsistent patterns would prompt them to reflect upon how their cycles may have varied across different stages of their lives. As Tuli et al. [54] emphasize, "an individual's association with menstrual tracking evolves with changing bodies and life choices across different transitions through their menstrual journey".

One of the main critiques of current menstrual trackers is that they often reinforce the false notion that there is a prototypical menstrual cycle consistent within and across individuals [18, 54]. For instance, when a menstrual tracking app assumes an idealized 28-day cycle as the standard, it can lead users to believe that deviations from this supposed norm are problematic. This can create anxiety or frustration among users whose cycles naturally vary, reinforcing an unrealistic expectation of uniformity.

Although most participants focused on the opportunities rendered by performing comparisons within their own data, some acknowledged the potential for expanding their menstrual health understanding by comparing their data with others **(G4)**. Participants envisioned that by comparing cycles side-by-side, they would be able to determine whether characteristics of their menstrual cycle fluctuated within healthy thresholds between cycles. Looking at other people's cycles could either perpetuate or challenge a "universal womanhood" experience [29] depending on the diversity of individuals and signal dimensions being presented. To avoid potential pitfalls, users should be provided with not only the ability to investigate the signal data from different cycles but also the context underlying that data. With this information, users may be able to reason about whether fair comparisons are being made or whether the data being examined represents the broad diversity of menstrual health experiences seen worldwide.

- **Refined Design Requirement:** The interface should enable users to compare various cycles, including their own historical cycles and those of their peers. *The interface should also explain the context behind cycle comparisons, whether they are based on different life stages within the same individual or individuals from diverse demographic groups or lifestyles* **(G3, G4)**.

6.6 Methodological Decisions and Future Work

Since current menstrual trackers do not allow users to incorporate data from external sources, we anticipated that they would have challenges envisioning how menstrual trackers could facilitate multivariate exploration. We presented medium-fidelity wireframes as design probes to solicit feedback while circumventing such technological limitations. Unlike traditional think-aloud methods, our

focus was not on evaluating participants' ability to navigate the interface successfully, but rather on their reactions to the presented concepts and the functional requirements that informed their design. We believe that goal-driven think-aloud evaluations like the one presented in this work could be used to probe important considerations in application domains beyond menstrual tracking.

In this study, we also chose not to present participants with real data. We adopted this approach because many participants had not collected the requisite data to populate our design probes, especially since we wanted them to consider the utility of seeing historical data from previous months. Additionally, not all participants used the same menstrual tracking apps or tracked the same data prior to enrollment. To standardize the presented concepts, we showed participants simulated data and asked them to project their responses to that data. Future work could involve deploying solutions in real-world settings to understand how our design requirements evolve over time with the usage of menstrual tracking interfaces. However, it is worth noting that some of the concepts that participants emphasized still require technological advancements before they can be feasibly deployed. For example, generating accurate symptom predictions **(DR5)** would involve developing new algorithms, while facilitating convenient support for multivariate data **(DR2)** would require a robust tool that can ingest data from heterogeneous sources. Accomplishing these goals would also require addressing privacy concerns and regulatory constraints surrounding the collection and usage of comprehensive personal health data, all of which are beyond the scope of our work.

Another limitation of this study is that our findings are predicated on the demographics of our population, which was skewed towards young adults who were not actively seeking pregnancy. While this population may represent a substantial portion of the primary user base for menstrual tracking apps [7], future research should explore how the findings from this study can be generalized to a broader range of age groups and users with diverse menstrual tracking needs and experiences (e.g., pregnancy and menopause). Finally, participants in this study did not extensively discuss the effort required to gather all of the data displayed. This may be due to the novelty and excitement around the prospect of accessing diverse multivariate data, although some participants did note that they may prioritize data collection based on perceived relevance and convenience. Future studies should explore how data collection behaviors influence people's interest, satisfaction, and perceived success in pursuing their menstrual-tracking goals.

7 CONCLUSION

Current menstrual trackers often lack the functionality to support users' menstrual-tracking goals beyond predictions and simplified symptom monitoring. We address these shortcomings by articulating and refining functional design requirements to support menstrual data exploration. Through a formative study and semi-structured interviews, we identified several key features that participants desired in future menstrual trackers: support for varied timeline representations for event planning and reflection, the ability to differentiate signals across menstrual cycle phases, the feature to group and prioritize signals based on perceived relevance and actionability, the capability to aggregate cycles across different time horizons for reflection, text-based explanations of underlying biological mechanisms, and the integration of contextual information for cycle comparison. Using these findings, we provide a refined list of functional design requirements to support users in achieving their goals in exploring menstrual health data.

ACKNOWLEDGMENTS

This research was funded in part by the Wolfond Scholarship in Wireless Information Technology, NSERC Discovery Grants #RGPIN-2021-03457, #RGPIN-2021-04268, and #RGPIN-2018-05072, and an unrestricted gift from Google.

REFERENCES

[1] Teresa Almeida, Rob Comber, and Madeline Balaam. 2016. HCI and Intimate Care as an Agenda for Change in Women's Health. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2599–2611.

[2] Teresa Almeida, Rob Comber, Gavin Wood, Dean Saraf, and Madeline Balaam. 2016. On looking at the vagina through Labella. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 1810–1821.

[3] Jeffrey Bardzell, Shaowen Bardzell, Amanda Lazar, and Norman Makoto Su. 2019. (Re-) Framing Menopause Experiences for HCI and Design. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[4] Jonathan Bidwell, Alan Leviton, Michele Jackson, and Ketan Mane. 2023. Sayantanee Das, Tobias Loddenkemper. The Need for Visualization Tools in the Electronic Health Record and in Decision Aids. *Journal of Bioinformatics and Systems Biology* 6 (2023), 183–186.

[5] Dagmar Celuchova Bosanska, Michal Huptych, and Lenka Lhotská. 2022. A Pipeline for Population and Analysis of Personal Health Knowledge Graphs (PHKGs). *CEUR Workshop Proceedings*.

[6] Tejinder Kaur Brar, K.D. Singh, and Avnish Kumar. 2015. Effect of different phases of menstrual cycle on heart rate variability (HRV). *Journal of Clinical and Diagnostic Research* (2015), CC01.

[7] Anna Broad, Rina Biswakarma, and Joyce C Harper. 2022. A survey of women's experiences of using period tracker applications: Attitudes, ovulation prediction and how the accuracy of the app in predicting period start dates affects their feelings and behaviours. *Women's Health* (2022), 17455057221095246.

[8] Rasheeta Chandler, Dominique Guillaume, Andrea Parker, Jessica Wells, and Natalie D Hernandez. 2021. Developing Culturally Tailored mHealth Tools to Address Sexual and Reproductive Health Outcomes Among Black and Latina Women: A Systematic Review. *Health Promotion Practice* (2021), 15248399211002831.

[9] Rasheeta Chandler, Dominique Guillaume, Andrea G Parker, Sierra Carter, and Natalie D Hernandez. 2020. Promoting Optimal Sexual and Reproductive Health with Mobile Health Tools for Black Women: Combining Technology, Culture and Context. *Perspectives on Sexual and Reproductive Health* 52, 4 (2020), 205–209.

[10] Mayara Costa Figueiredo, Thu Huynh, Anna Takei, Daniel A Epstein, and Yunan Chen. 2021. Goals, life events, and transitions: examining fertility apps for holistic health tracking. *JAMIA Open* (2021), ooab013.

[11] Sara Di Bartolomeo, Aditeya Pandey, Aristotelis Leventidis, David Saffo, Uzma Haque Syeda, Elin Carstensdottir, Magy Seif El-Nasr, Michelle A Borkin, and Cody Dunne. 2020. Evaluating the effect of timeline shape on visualization task performance. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[12] Helen S Driver and Fiona C Baker. 1998. Menstrual factors in sleep. *Sleep Medicine Reviews* 2, 4 (1998), 213–229.

[13] Anna Eaglin and Shaowen Bardzell. 2011. Sex toys and designing for sexual wellness. In *Extended Abstracts of the 2011 CHI Conference on Human Factors in Computing Systems*. 1837–1842.

[14] Daniel A Epstein, Nicole B Lee, Jennifer H Kang, Elena Agapie, Jessica Schroeder, Laura R Pina, James Fogarty, Julie A Kientz, and Sean Munson. 2017. Examining menstrual tracking to inform the design of personal informatics tools. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 6876–6888.

[15] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A Lived Informatics Model of Personal Informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 731–742.

[16] Jordan Eschler, Amanda Menking, Sarah Fox, and Uba Backonja. 2019. Defining Menstrual Literacy With the Aim of Evaluating Mobile Menstrual Tracking Applications. *CIN: Computers, Informatics, Nursing* (2019), 638–646.

[17] Margaret Flemings, Shanzay Kazmi, Rachel Pak, and Orit Shaer. 2018. Crimson Wave: Shedding Light on Menstrual Health. In *TEI '18: Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, 343–348.

[18] Sarah Fox and Daniel A Epstein. 2020. Monitoring Menses: Design-Based Investigations of Menstrual Tracking Applications. In *The Palgrave Handbook of Critical Menstruation Studies*. Palgrave Macmillan, Singapore, 733–750.

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

[20] Katie Gambier-Ross, David J McLernon, and Heather M Morgan. 2018. A mixed methods exploratory study of women's relationships with and uses of fertility tracking apps. *Digit Health* (2018), 2055207618785077.

[21] Olivier Garot, Julian Rössler, Juliane Pfarr, Michael T Ganter, Donat R Spahn, Christoph B Nöthiger, and David W Tscholl. 2020. Avatar-based versus conventional vital sign display in a central monitor for monitoring multiple patients: a multicenter computer-based laboratory study. *BMC medical informatics and decision making* 20, 1 (2020), 1–10.

[22] Greta Gasciauskaite, Justyna Lunkiewicz, Tadzio R Roche, Donat R Spahn, Christoph B Nöthiger, and David W Tscholl. 2023. Human-centered visualization technologies for patient monitoring are the future: a narrative review. *Critical Care* 27, 1 (2023), 254.

[23] Brianna Mae Goodale, Mohaned Shilaih, Lisa Falco, Franziska Dammeier, Györgyi Hamvas, and Brigitte Leeners. 2019. Wearable sensors reveal menses-driven changes in physiology and enable prediction of the fertile window: observational study. *Journal of medical Internet research* 21, 4 (2019), e13404.

[24] Joseph Henrich, Steven J Heine, and Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and brain sciences* 33, 2-3 (2010), 61–83.

[25] Sarah Homewood. 2019. Inaction as a design decision: Reflections on not designing self-tracking tools for menopause. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[26] Sarah Homewood and Anna Vallgårda. 2020. Putting Phenomenological Theories to Work in the Design of Self-Tracking Technologies. In *Proceedings of the 2020 Designing Interactive Systems Conference*. ACM, 1833–1846.

[27] Dandan Huang, Melanie Tory, and Lyn Bartram. 2016. A Field Study of On-Calendar Visualizations. *Proceedings - Graphics Interface* (2016).

[28] Nassim Jafarinaimi, Jodi Forlizzi, Amy Hurst, and John Zimmerman. 2005. Breakaway: an ambient display designed to change human behavior. In *Extended Abstracts of the 2005 CHI Conference on Human Factors in Computing Systems*. 1945–1948.

[29] Os Keyes, Burren Peil, Rua M Williams, and Katta Spiel. 2020. Reimagining (women's) health: HCI, gender and essentialised embodiment. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 4 (2020), 1–42.

[30] Sung-Hee Kim. 2022. A Systematic Review on Visualizations for Self-Generated Health Data for Daily Activities. *International Journal of Environmental Research and Public Health* 19, 18 (2022). https://doi.org/10.3390/ijerph191811166

[31] Neha Kumar, Naveena Karusala, Azra Ismail, and Anupriya Tuli. 2020. Taking the Long, Holistic, and Intersectional View to Women's Wellbeing. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 4 (2020), 1–32.

[32] Amanda Lazar, Christian Koehler, Joshua Tanenbaum, and David H. Nguyen. 2015. Why we use and abandon Smart Devices. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (2015).

[33] Amanda Lazar, Norman Makoto Su, Jeffrey Bardzell, and Shaowen Bardzell. 2019. Parting the red sea: Sociotechnical systems and lived experiences of menopause. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–16.

[34] Johanna Levy and Nuria Romo-Avilés. 2019. "A good little tool to get to know yourself a bit better": a qualitative study on users' experiences of app-supported menstrual tracking in Europe. *BMC Public Health* 19, 7 (2019).

[35] Georgianna E Lin, Elizabeth D Mynatt, and Neha Kumar. 2022. Investigating Culturally Responsive Design for Menstrual Tracking and Sharing Practices Among Individuals with Minimal Sexual Education. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.

[36] Deborah Lupton. 2014. Self-tracking modes: Reflexive self-monitoring and data practices. *Available at SSRN 2483549* (2014).

[37] Anna Maijala, Hannu Kinnunen, Heli Koskimäki, Timo Jämsä, and Maarit Kangas. 2019. Nocturnal finger skin temperature in menstrual cycle tracking: ambulatory pilot study using a wearable Oura ring. *BMC Women's Health* 19, 1 (2019), 1–10.

[38] J. Maitland, S. Sherwood, L. Barkhuus, I. Anderson, M. Hall, B. Brown, M. Chalmers, and H. Muller. 2006. Increasing the Awareness of Daily Activity Levels with Pervasive Computing. In *2006 Pervasive Health Conference and Workshops*. 1–9.

[39] Alessandro Marcengo and Amon Rapp. 2014. Visualization of Human Behavior Data: The Quantified Self. In *Innovative Approaches of Data Visualization and Visual Analytics*. IGI Global, 30.

[40] Sonali R. Mishra, Predrag Klasnja, John MacDuffie Woodburn, Eric B. Hekler, Larsson Omberg, Michael Kellen, and Lara Mangravite. 2019. Supporting Coping with Parkinson's Disease Through Self Tracking. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–16.

[41] Guilherme Oliveira, João Comba, Rafael Torchelsen, Maristela Padilha, and Claudio Silva. 2013. Visualizing running races through the multivariate time-series of multiple runners. In *2013 XXVI Conference on Graphics, Patterns and Images*. IEEE, 99–106.

[42] Juliane Pfarr, Michael T Ganter, Donat R Spahn, Christoph B Noethiger, and David W Tscholl. 2020. Effects of a standardized distraction on caregivers' perceptive performance with avatar-based and conventional patient monitoring: A multicenter comparative study. *Journal of Clinical Monitoring and Computing* 34 (2020), 1369–1378.

[43] Adrienne Pichon, Kasey B Jackman, Inga T Winkler, Chris Bobel, and Noémie Elhadad. 2022. The messiness of the menstruator: assessing personas and functionalities of menstrual tracking apps. *Journal of the American Medical Informatics Association* (2022), 385–399.

[44] Nidhi Rastogi and Mohammed J Zaki. 2020. Personal health knowledge graphs for patients. *arXiv preprint arXiv:2004.00071* (2020).

[45] Tadzio R Roche, Sadiq Said, Julia Braun, Elise JC Maas, Carl Machado, Bastian Grande, Michaela Kolbe, Donat R Spahn, Christoph B Nöthiger, and David W Tscholl. 2021. Avatar-based patient monitoring in critical anaesthesia events: a randomised high-fidelity simulation study. *British journal of anaesthesia* 126, 5 (2021), 1046–1054.

[46] Minna Ruckenstein. 2014. Visualized and interacted life: Personal analytics and engagements with data doubles. *Societies* 4, 1 (2014), 68–84.

[47] Young June Sah and Wei Peng. 2015. Effects of visual and linguistic anthropomorphic cues on social perception, self-awareness, and information disclosure in a health website. *Computers in Human Behavior* 45 (2015), 392–401.

[48] Jessica Schroeder, Chia-Fang Chung, Daniel A. Epstein, Ravi Karkar, Adele Parsons, Natalia Murinova, James Fogarty, and Sean A. Munson. 2018. Examining Self-Tracking by People with Migraine: Goals, Needs, and Opportunities in a Chronic Health Condition. In *Proceedings of the 2018 Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 135–148.

[49] Moushumi Sharmin, Andrew Raij, David Epstien, Inbal Nahum-Shani, J Gayle Beck, Sudip Vhaduri, Kenzie Preston, and Santosh Kumar. 2015. Visualization of time-series sensor data to inform the design of just-in-time adaptive stress interventions. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 505–516.

[50] Mohaned Shilaih, Valérie de Clerck, Lisa Falco, Florian Kübler, and Brigitte Leeners. 2017. Pulse rate measurement during sleep using wearable sensors, and its correlation with the menstrual cycle phases, a prospective observational study. *Scientific Reports* 7, 1 (2017), 1–7.

[51] Barbara Sommer. 1978. Stress and menstrual distress. *Journal of Human Stress* 4, 3 (1978), 5–47.

[52] Margarent L. Stubbs and Evelina W. Sterling. 2020. Learning About What's "Down There": Body Image Below the Belt and Menstrual Education. In *The Palgrave Handbook of Critical Menstruation Studies*. Palgrave Macmillan, Singapore, 233–252.

[53] Da Tao, Juan Yuan, Xingda Qu, Tieyan Wang, and Xingyu Chen. 2018. Presentation of personal health information for consumers: An experimental comparison of four visualization formats. *Engineering Psychology and Cognitive Ergonomics* (2018), 490–500.

[54] Anupriya Tuli, Surbhi Singh, Rikita Narula, Neha Kumar, and Pushpendra Singh. 2022. Rethinking Menstrual Trackers Towards Period-Positive Ecologies. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–20.

[55] Agatha Tutia, Kelda Baljon, Lan Vu, and Daniela K Rosner. 2019. HCI and Menopause: Designing with and around the Aging Body. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–8.

[56] Doreen J. Wetli, Lisa Bergauer, Christoph B. Nöthiger, Tadzio R. Roche, Donat R. Spahn, David W. Tscholl, and Sadiq Said. 2022. Improving Visual-Patient-Avatar Design Prior to Its Clinical Release: A Mixed Qualitative and Quantitative Study. *Diagnostics* 12, 2 (2022).

[57] Alyson L. Young and Andrew D. Miller. 2019. "This Girl is on Fire": Sensemaking in an Online Health Community for Vulvodynia. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[58] Rhonda Zwingerman, Michael Chaikof, and Claire Jones. 2020. A Critical Appraisal of Fertility and Menstrual Tracking Apps for the iPhone. *Journal of Obstetrics and Gynaecology Canada* 42, 5 (2020), 583–590.