Engaging research participants with self-logged menstrual health data

SAMANTHA ROBERTSON, University of California, Berkeley, USA
KIM HARLEY, University of California, Berkeley, USA
NILOUFAR SALEHI, University of California, Berkeley, USA

Fig. 1. It is increasingly common for researchers to use data from personal health tracking apps in menstrual health research, but this work rarely engages directly with the users who contributed the data. Our work explores how adding research participants into the data analysis loop could help users understand more about their health, and improve the quality of research conducted with this data.

Many people use health tracking apps to keep track of their menstrual cycles, often in the hopes of better understanding their own health, and being able to identify when something might be wrong. However, it can be very difficult to interpret this data alone. Meanwhile, it is becoming increasingly common for researchers to use data from these apps to learn more about menstrual health. In this work we ask, how could more participatory approaches to conducting menstrual health research benefit both participants and researchers? We identify key challenges and risks of this kind of engagement, and propose four design guidelines for human-in-the-loop data analysis tools that engage participants with large-scale, quantitative menstrual health research: surface and elicit feedback on the data cleaning and analysis procedure; convey information relative to other users and clinical guidance; structure engagement to ensure valid analyses; and support social engagement and learning. For each of these, we highlight key open research questions relevant to the HILDA and visualization research communities. We plan to evaluate and iterate on these guidelines through design workshops with users, researchers, and healthcare providers.

Additional Key Words and Phrases: personal health informatics, data visualization, participatory research

ACM Reference Format:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2022 Association for Computing Machinery.
Manuscript submitted to ACM
1 INTRODUCTION

As more people have started using technology for tracking their health and habits [23, 30], researchers have become interested in using that data for scientific research. For example, with the growing popularity of apps for tracking menstrual cycles [3, 26], this has become a popular approach in menstrual health research [32, 58, 72, 76]. The findings of this research could help users of menstrual tracking apps improve their understanding of their own cycles relative to others’, and learn more about menstrual health more broadly [11, 15, 16, 20, 27, 45, 53, 74, 85]. Users could also provide useful insights and feedback on the research, drawing on their lived expertise and deep contextual knowledge of their data [54, 73]. However, most of this research is not currently conducted in a way that fosters bidirectional learning between researchers and participants (Fig 1). Our work seeks to understand whether, and if so what kind of, interactive data tools could foster this kind of engagement in a way that maximizes benefits to users, and minimizes risks. In this paper we seek to define and motivate a research agenda to develop more collaborative, participatory, and accessible methods for conducting quantitative research with personal data.

In this paper we focus on menstrual tracking and menstrual health research as a case study. People track their menstrual cycles to build general self-awareness [26, 30], to manage and plan around their periods [26, 52] or symptoms of specific conditions like fibroids, endometriosis, or bleeding disorders [11, 52], to conceive or avoid pregnancy [15, 26, 30, 53], and to learn about menstrual health generally. However, it is difficult for users to derive meaningful and interesting insights from their data with existing tools, especially without clinical guidance [10, 47] or the ability to compare their experiences to others’ [16]. Although researchers have developed a wide variety of bespoke visualizations to engage people with self-tracking data [4, 5, 13, 22, 24, 28, 29, 34, 35, 38–43, 48, 51, 63, 68–71], most have focused on individuals engaging with their own data, often in domains like wellness and productivity, where users’ goals are more often oriented towards behavior change, and the public has greater general knowledge about the domain [23, 26]. In this work we explore how interactive data analysis tools could be used to engage users with menstrual health research conducted with data from period tracking apps.

In an ongoing research project, we are using self-logged menstrual cycle data from over 6,000 users of a menstrual tracking app, Clue by BioWink GmbH, to understand how menstrual cycle characteristics in adolescents are associated with factors like stress, and sleep. This research presents an opportunity to actively engage participants in analyzing their own data, and comparing it to the study population. In this paper, we explore the possible benefits of increasing participation in research with menstrual tracking data, as well as potential risks and challenges. We propose four design guidelines for systems that could support this kind of engagement: surface and elicit feedback on the data cleaning and analysis procedure; convey information relative to other users and clinical guidance; structure engagement to ensure valid analyses; and support social engagement and learning. We plan to test and iterate on these guidelines through design workshops with users, researchers, and healthcare providers.

Engaging users who may have limited knowledge of both menstrual health and data analysis in collaborative engagement with researchers around menstrual health data opens up new directions for research into human-in-the-loop data analysis tools. For instance, how can we engage a large group of people in analyzing, sharing, and co-earning with and through their collective data? How can we help people interpret their data in the context of a broader sample in a way that accounts for uncertainty and the wide range of “normal” menstrual cycle characteristics? How can we incorporate expert guidance into the data exploration experience without being overly prescriptive about how users can explore their data? We look forward to feedback from the HILDA community about how we can adapt the tools that have been developed in this field to this new domain of menstrual health.
2 ENGAGING PARTICIPANTS IN RESEARCH WITH PERSONAL DATA: BENEFITS AND CHALLENGES

Across disciplines, researchers have worked to foster collaboration with participants in their studies. For example, the participatory research and citizen science methodologies center participants as key contributors to the work, from formulating research questions, to collecting and interpreting data [21, 37]. In this section we draw on literature from these traditions to motivate our work and anticipate risks and challenges.

2.1 Benefits

Engaging with menstrual health research using tracking data could help participants learn more about their own health & menstrual health more broadly, a key goal for many people who track their periods [11, 15, 16, 27, 45, 52, 53, 74, 85]. Both the participatory research and citizen science traditions emphasize the importance of sharing findings with those who contributed to the research, with participants’ learning being a key benefit of this practice [2, 6, 7, 75]. Prior work has found that tracking app users are interested in comparing their data to others’, and even share their data online when comparison is not supported by existing tools [16, 81]. Allowing users to explore their own data relative to a research study cohort could support these kinds of comparisons and help users interpret their data. Engaging users in research also offers opportunities for them to learn about the research process and improve their data literacy [44, 75].

Participatory research not only offers educational opportunities for participants, but also enables researchers to draw on participants’ lived experiences and contextual expertise. External context is invaluable to the interpretation of personal data [54, 73]. For example, a participant in Moore et al.’s study of home air quality data associated large spikes in poor air quality with times they had cooked bacon – an insight that researchers working with this data could never have derived alone [54]. People’s lived experiences and contextual understanding of their data could improve many stages of the research process, from forming relevant and previously overlooked hypotheses [52], to more accurate data cleaning and aggregation [54]. For example, a major challenge in the context of menstrual tracking data is differentiating very long cycles from lapses in app usage. Users may be able to provide much more accurate information about their cycles and app usage than could be inferred from the data alone.

Finally, our work is an opportunity to increase participant awareness and oversight of large-scale quantitative research. A 2014 study conducted as part of the Health Data Exploration Project found that people are open to sharing their personal health data with researchers, as long as they know it will not be used for commercial purposes and that the research will have some personal or public benefit [1]. McKillop et al. conducted workshops with women with endometriosis, and also found that they were open to participating in research involving self-logged personal data, especially if the research would help younger people with the condition [52]. Unfortunately, research with large-scale personal data has not always met participants’ expectations [33, 65]. This has led to calls for ensuring that data subjects are aware of research conducted with their data, and have some power to shape the kinds of questions asked and how the benefits of the research are distributed [59, 65].

2.2 Risks and Challenges

Achieving these benefits in practice will be challenging, especially at the scale of many quantitative research studies, which can involve thousands, or even millions, of participants distributed around the world [58].

An important challenge will be preventing misinterpretation, over-diagnosis, and unnecessary concern for users. One of the major issues with existing personal health technologies is a lack of clinical validation [19, 27, 87].
Period tracking apps have had to withdraw even clinically validated diagnostic screening tools after they raised concern about over-diagnosis and unnecessary stress for users [66]. In the context of fertility tracking, research has found that aggregating data across users can help individuals better interpret their own data, but it can also enforce norms and generate stress and anxiety, especially when it is unclear how clinically valid the aggregate analyses are [16]. This presents a difficult trade-off between empowering people to better understand their health, and exercising caution to avoid misinformation and unnecessary stress. Researchers across the health sciences have grappled with this challenge and put forth best practices, such as providing personalized support and clinical guidelines to help people interpret study results [9, 55], but these practices remain contested [55]. One goal of this work is to better understand these risks and develop mitigation strategies in the context of participant engagement with menstrual health research with tracking data.

Another critical challenge will be protecting participants’ privacy. Menstrual cycles and other personal health data are extremely personal and intimate data sources, and users’ privacy must be a top priority. Tools for engaging with research data must be carefully designed to protect individuals’ and vulnerable subpopulations’ information. Unfortunately, many period tracking apps have violated users’ trust by selling their data to third parties [1]. Although research suggests that many people are open to sharing data with researchers, it is not clear how many users know how their data is used, and by whom, nor do we have a clear understanding of tracking app users’ contextual expectations of privacy [1]. A first step towards this understanding is ensuring that participants are aware of how their data is being used, and give them more opportunities to opt out of participating. We believe that engaging users with this kind of research more actively will be a useful way to begin developing norms and expectations around privacy, and empower users to advocate for their privacy and make informed decisions about which technologies to use. Of course, a possible outcome of increased user awareness of data sharing with researchers is restricted access to data for researchers; our goal is to better understand how engaging users with research might benefit both users and researchers, but we recognize that a possible outcome of this work is finding that the risks of such engagement may outweigh the benefits.

### 3 DESIGN GUIDELINES

Engaging users and researchers with personal health data in a mutually beneficial way will require interactive systems that allow users to explore their data with neither technical nor domain expertise. Research in the human-in-the-loop data analytics (HILDA) and broader data visualization community has explored ways of making it easier for a range of audiences to engage with large-scale quantitative data. For example, interfaces that allow users to build visualizations using natural language [31, 62, 64], or demonstrations (e.g., other visualizations, sketches, or constraints) [60, 61, 67, 79, 80, 83, 84], or that recommend relevant visualizations based on partial specifications [50, 56, 82] or interesting patterns in a dataset [38, 39, 46] all have great potential for allowing end-users of self-tracking apps to freely explore and visualize their data without needing to code.

However, may existing tools for analyzing and visualizing data without needing to write code have been designed for domain experts. People who use tracking apps may have deep expertise on their own experiences, but many will have limited menstrual health literacy and data literacy. In this section, we discuss four design guidelines for systems to engage these people with research data, informed by prior work on data visualization for personal health data and participatory research. We highlight open research questions relevant to the HILDA community that each of these guidelines raise. Figure 2 envisions how these guidelines might be applied to visualization tools, using menstrual cycle length and regularity as an example. These guidelines are hypotheses about how we might be able to support engagement, maximizing the benefits and mitigating the risks outlined in the previous section. In the following section
Fig. 2

we will discuss how we plan to test and refine these guidelines in the context of an ongoing menstrual health research project.

**DG1** Surface and elicit feedback on the data cleaning and analysis procedure. How data has been collected, cleaned, and aggregated is often abstracted away in data visualization systems, but it is also an extremely consequential process involving many value-laden decisions [8, 14]. Surfacing this process is key to giving users meaningful oversight of the research, as well as leveraging their expertise to improve the effectiveness of data cleaning. For instance, as mentioned above, a particularly important case for period tracking data is removing cycles where a user stopped using or forgot to use the app. However, it is difficult to communicate complex analysis methods in a way that is digestible and understandable to a broad audience [8, 14]. Providing editing tools for users to annotate these cycles not only gives researchers the opportunity to improve their data cleaning, but also to learn more about the data and the participants. For example, users could be asked follow-up survey questions about why they are editing their data (Fig. 2(a)).

**Open questions:** How can tools surface the data cleaning and analysis methods in a way that is understandable and useful to participants? How can we elicit feedback from users in a way that can be integrated back into the analysis procedure? How can we ensure that users understand how their feedback has been accounted for?

**DG2** Convey uncertainty and variability when making comparisons. Comparison against baselines and peers is important for users to be able to contextualize and interpret their data [9, 16, 36, 38] (Fig. 2(b)). However, comparison can also promote strict norms and create stress for those who appear to fall outside of them [16]. Tools should provide information to help people interpret these comparisons, mitigate unnecessary stress, and connect users to healthcare providers or other sources of support [9, 55]. It is an open question how visualizations can or should convey bad news and have an appropriate impact on an audience [9, 14], and this is likely to be even more challenging when a visualization is conveying information about the viewer’s own health, since people engage most with visualizations that they can directly relate to [57]. This issue is complicated by the fact that menstrual cycles are extremely variable and individual, so comparative visualizations may often look concerning to a user when they have nothing to worry about. On the other hand, many people who do have health conditions related to their menstrual cycles, e.g. endometriosis, struggle for years to get a diagnosis [16]. There is a balance to strike between avoiding unnecessary concern and over-diagnosis, while also not dismissing people’s valid health concerns [11].

**Open questions:** How should we visualize uncertainty in the context of highly variable and individualized distributions? Can uncertainty visualization mitigate unnecessary stress and over-diagnosis? What are the implications of allowing comparative analyses for individual and
sub-group privacy? Could we automatically detect when someone views a visualization that might be worrisome and offer relevant resources?

(DG3) Structure engagement to ensure valid analyses. Ideally, participants should be able to explore the data flexibly and answer questions that are most personally relevant to them. Tools that allow users to create visualizations using direct manipulation and/or natural language [31, 60–62, 64, 67, 79, 80, 83, 84], or even visualization recommendation systems [38, 39, 46, 50, 56, 82], are likely to be important to support creative and open-ended engagement. A challenge with supporting open-ended exploration, however, will be ensuring that users do not conduct meaningless or misleading analyses [14, 86]. Identifying questions that can be answered with a given dataset is challenging even for experienced analysts [17]. Engagement should therefore allow open-ended engagement as much as possible while remaining structured enough to guide users to valid analyses and interpretations. For example, users should be guided towards reasonable comparisons and baselines, and away from trying to make causal inferences from observational data. Open questions: Can we automatically detect when a user tries to conduct an analysis that is not well supported by the data? How can we interactively teach users how to ask valid questions and interpret their results?

(DG4) Support social engagement and learning. Research on personal health informatics is increasingly acknowledging the social dimensions of health tracking [12, 25, 49, 77]. Social engagement, even if passive, is important for learning [18, 74, 78] and can be a source of support, particularly for people with chronic health conditions [11, 45]. Much of this kind of support and collective sense-making defies quantification and analysis in the way that would be supported by engaging people on an isolated, individual basis with quantitative menstrual health research. Open questions: How can we support asynchronous and (possibly) anonymous collaboration in understanding and contributing to quantitative health research? How does social engagement mitigate or worsen the risks of misinformation and over-diagnosis?

4 FUTURE WORK

The guidelines we have proposed are hypotheses informed by prior literature, and will need to be tested and validated before being more widely used in the field. The first two authors are currently using data from Clue by BioWink GmbH, a popular period tracking app, to characterize menstrual cycles in adolescents, and investigate how cycle characteristics are associated with behavioral factors like sleep and stress. Our future work will involve engaging participants with this research more deeply, evaluating and adapting these guidelines in the process.

In the longer term, we hope to see research conducted in collaboration with participants from generating research questions and hypotheses, through to dissemination. We believe the HILDA community can play a critical role in building tools to support this vision.

5 CONCLUSION

In this paper we argued that engaging participants directly in research with large-scale data from menstrual cycle tracking apps has great potential to benefit participants and improve research. However, this engagement will not be without risk, especially around misinterpretation and over-diagnosis, and privacy. We presented early design guidelines for systems that could support this engagement, and described our plans to evaluate and iterate on them through design workshops. We look forward to feedback from the HILDA community about how we can extend tools for exploring and visualizing data without code, as well as how to visualize individuals’ data in the context of an uncertain and highly variable distribution.
ACKNOWLEDGMENTS

We are indebted to the more than 10,000 people who participated in our survey and contributed their tracking data to our research, whose contributions inspired this work. We would also like to thank Amanda Shea and Charlie Upton from Clue by BioWink GmbH for their ongoing support. We are grateful to Leilani Battle and the members of the EPIC lab at U.C. Berkeley for their helpful feedback on this work.

REFERENCES


