

Harnessing the Power of Patient-Generated Data

Eun Kyoung Choe
University of Maryland,
College Park

Bongshin Lee
Microsoft Research

Tariq Osman Andersen
University of Copenhagen

Lauren Wilcox
Georgia Institute of
Technology

Geraldine Fitzpatrick
TU Wien

Department editor:
Sarah Clinch;
sarah.clinch@manchester.ac.
uk

The authors report on the PervasiveHealth 2017 workshop, “Leveraging Patient-Generated Data (PGD) for Collaborative Decision Making in Healthcare.” They discuss characteristics of PGD, followed by scenarios demonstrating the data-sharing practice among patients, clinicians, and caregivers. The authors also highlight current challenges and opportunities, and outline a future research agenda to envision ways to harness the power of PGD.

We are living in an exciting time that is experiencing a paradigm shift in health and medical sciences. In recent years, personal health technologies have emerged that allow patients to collect a wide range of health-related data outside the clinic. These patient-generated data (PGD) reflect patients’ everyday behaviors including physical activity, mood, diet, sleep, and symptoms. Thus, sharing PGD between patients and clinicians can help them communicate about health-related concerns and identify actionable insights. Despite active research in this area, however, it is still unclear how we should go about leveraging these PGD to integrate them into clinical practice.

On 23 May 2017, a group of 16 researchers working at the intersection of health informatics, human-computer interaction, and visualization gathered to discuss challenges and opportunities in leveraging PGD in healthcare. The workshop organizers had three objectives:

On 23 May 2017, a group of 16 researchers working at the intersection of health informatics, human-computer interaction, and visualization gathered to discuss challenges and opportunities in leveraging PGD in healthcare. The workshop organizers had three objectives:

- identify contexts where PGD sharing could be useful for patients and clinicians,
- bridge communities working in related fields, and
- distill a set of lessons learned from creating scenarios and working toward solutions as a starting point for future efforts.

The workshop was part of 11th European Alliance for Innovation (EAI) International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth 2017).

When planning the workshop, the organizers quickly agreed that looking for a one-size-fits-all solution would be misguided, because the design space of PGD is too broad and complex. Thus, they asked the participants to describe in their position paper a concrete and specific PGD-

sharing context they were working on. At the beginning of the workshop, participants presented their PGD research contexts and key challenges in their work. Based on this shared understanding, they collectively delineated a set of dimensions to better characterize PGD, and designed example scenarios considering stakeholders' contexts and goals. In this article, we report back key themes, scenarios, and an underexplored research agenda in PGD sharing contexts that emerged from the workshop discussion.

CHARACTERISTICS OF PATIENT-GENERATED DATA (PGD)

PGD cover a wide range of health-related data that are created, recorded, and gathered by patients, family members, or other caregivers.¹ PGD overlap with patient-reported outcomes (PROs), especially when the PROs reflect patients' health conditions (as opposed to their satisfaction with a healthcare entity, for instance). As PGD data types are diverse, so too are the situations and contexts in which these data can be shared and the approaches to enabling data sharing. Here we describe four PGD dimensions and discuss how they shape specific PGD sharing contexts.

First, *who initiates the tracking*—whether it is a patient, clinician, caregiver, or healthcare entity—influences all aspects of the sharing process: from motivation for collecting data, to the practice of collecting and utilizing the data.² If clinicians can guide the patient's tracking throughout the care practice to help ensure data quality, this could increase clinicians' willingness to trust and rely on the data in the long run.

Second, *tracking purpose* necessarily interacts with the first dimension. Researchers have recently identified various reasons why people track personal data.³ For example, people track data to reflect on past behaviors, be aware of their behaviors, solve a particular problem, or fulfill their curiosity. Additionally, clinicians have different purposes for wanting their patients to track data. For example, they can use PGD to better understand what behaviors are “typical” for the patient and to augment communication in the consultation,⁴ as well as to aid in the diagnosis of the patient's condition⁵; they also see tracking as a way to improve patient engagement and health outcomes.⁶ Hospitals might use PGD instrumentally for quality monitoring, which may be less useful or visible for patients.

Third, in research contexts, we do not often address *data storage and ownership* issues because the focus of research is often to show a proof of concept. In clinical practice, however, where PGD get stored (whether in a hospital-owned electronic medical record, commercial platform, or patient's notebook) and who directly owns the data affect the practical feasibility and willingness of clinicians to access and leverage PGD.

Finally, different *capture mechanisms* exist. For example, sleep quality can be measured in a subjective manner through self-reporting as well as in an objective and automated manner using a tracking device (for example, restless sleep in minutes). The complementary data from multiple sources allow people to cross-check the reliability of the data and identify insights, although the process of analyzing and comparing the data might be difficult and confusing.

PGD-SHARING SCENARIOS

Scenarios help us understand and analyze how technology could be used to reshape users' activities before a system is built.⁷ During the breakout sessions, workshop participants designed three PGD-sharing scenarios, considering a range of contexts and goals in addition to the PGD characteristics described in the previous section. People might have different health conditions, which can in turn lead to different goals: patients with chronic illnesses have different needs than acute patients and healthy individuals. Patients who already have a diagnosis manage clearly defined problems, whereas people with undiagnosed or pre-diagnosed conditions often deal with ambiguous and exploratory problems. Their technology proficiency, age, education level, income level, and access to care could widely vary as well. In the following three scenarios, we describe how

tracking personal data and sharing the data with clinicians might help them address some of the challenges within each scenario.

Scenario 1: Chronic Patient Care of an Older Adult with Low Motivation and Low Technology Literacy

We created this scenario to illustrate an extremely challenging type of patient—not only for clinicians but also for researchers. Sarah is a 68-year-old woman who was diagnosed with type 2 diabetes 15 years ago. She lives alone in a low-income residence located in a suburban city in the US. Although she has access to healthcare, Sarah has limited insurance coverage and worries about her medical bill. She suffers from mild depression and fears the possibility of developing dementia as it runs in her family. Sarah is not familiar with high-tech devices and does not own a smartphone, though she does have Internet access. Her lifestyle is in general sedentary and she has low motivation to manage her health (see Figure 1).

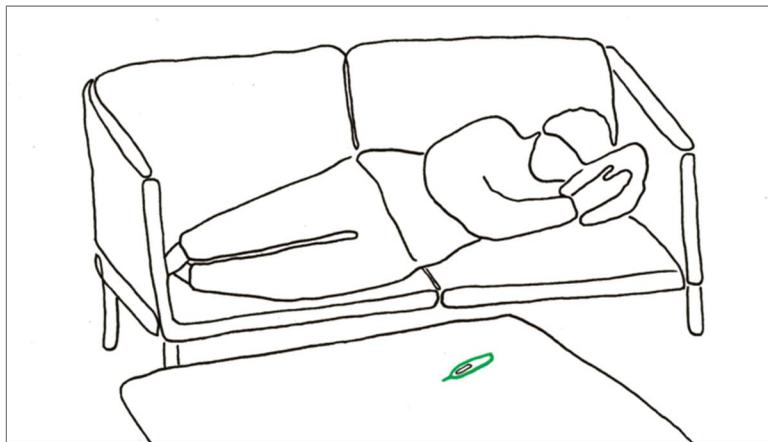


Figure 1. Scenario 1—Sarah, a diabetes patient. Her clinician wants to introduce self-tracking to Sarah as a way to engage her in the care practice.

Sarah dreads experiencing diabetes symptoms such as extreme fatigue and blurry vision as well as more serious complications such as foot problems and kidney failure. Her goal is to avoid these symptoms and complications of diabetes, but it is not clear to her how daily exercise and blood-glucose management connect to these outcomes.

In this scenario, Sarah's goals are not fully aligned with those of her clinician: the clinician wants her to exercise regularly and eat a healthy diet for long-term management, whereas Sarah only wants to avoid the immediate symptoms and complications of her illness. In this case, “low-tech” self-tracking (initiated by the clinician) and data sharing could reduce the gap between the clinician's goals and Sarah's goals. The clinician can introduce self-tracking to Sarah by asking her to start minimal tracking of her exercise and diabetes symptoms; she can enter her duration of exercise (for example, walks) and occurrence of symptoms via a text message or through a tracking website.

They can then use the self-tracking data to discuss the relationship between her day-to-day activities and her diabetes symptoms. In addition, to promote and sustain Sarah's engagement, the clinician can use the data to communicate the importance of tracking in their mutual understanding of Sarah's condition and progress. As Sarah gets used to the tracking, the clinician can also guide her to revise the tracking items. Together they can identify the items that are more important to collect and share, then focus on them; potential data to share include blood-sugar levels, medication taking, diet, mood, stress, and physical activity.

Scenario 2: Casual Data Tracking for Later Use in a Clinical Setting

This scenario is meant to illustrate the potential value of incidental or unpurposeful tracking of “life data” and how they may later become useful for health purposes. Oliver is in his mid-thirties and runs his own hair salon in a large European city. He is trained as a professional hairdresser and has experience from various high-end establishments. Besides his passion for healthy diets and fashion, he finds technology fascinating. Inspired by a close friend, Oliver was one of the first adopters of heart-rate monitors in the 1990s, and ever since he has been exploring a broad range of activity tracking including heart rate, sleep, food, and mood as well as more unconventional measures like gut bacteria and social activity (see Figure 2).

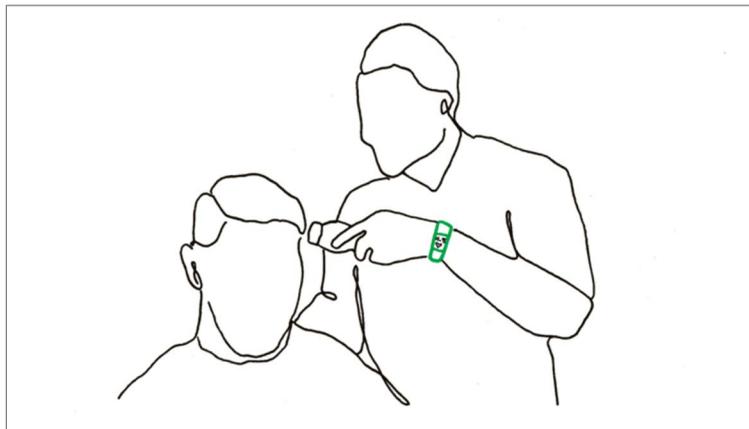


Figure 2. Scenario 2—Oliver, a casual self-tracker. Oliver and his clinician want to utilize Oliver’s self-tracking data for a better diagnosis.

His business was going well for a while, but lately sales have started to drop—mainly because of the bad temper and negative attitude Oliver shows toward his customers. Oliver’s social life has also changed. His close friends no longer call or write to him and he feels alone most of the time. Left in a life crisis and feeling more and more anxious, he makes an appointment with his doctor. At the meeting, they discuss his social history, how he sometimes loses his temper, and the times he feels anxious and emotionally unstable. Oliver also shows several printouts of some of the PGD he has been tracking over many years.

The doctor cannot reach a diagnosis and suggests that Oliver consult a colleague—a health “informatician” who is skilled in health data analyses and provides services in making sense of large-scale PGD. Upon specific requests, Oliver collects a multitude of different data including self-tracked data from devices and other behavioral data from sources like Facebook, and uploads them to the informatician. At the following consultation, the doctor shows Oliver different examples of data correlations: high blood pressure during periods of low social interaction, high levels of blood glucose associated with bad mood, and increased use of mobile applications correlated with low concentration. Based on the data analysis, the doctor diagnoses Oliver with a mild form of ADHD and begins treatment to improve and restore Oliver’s life.

Scenario 3: Parents with a Newborn Baby Capturing and Sharing Health and Developmental Data

This scenario covers a case in which individuals are not capable of tracking their data on their own, and thus caregivers need to capture and share the data on their behalf. Jessica and John are in their early 30s, with a son, Lucas, born just two weeks ago. They both work for IT companies in the US and have good health benefits.

As first-time parents, they have a lot to learn and are keen on keeping a close eye on the baby's health and development. Jessica and John want to figure out basic things including how to feed him and train him for sleep. They are also anxious to know whether the baby is growing according to standard milestones or having any issues. Thus, in addition to logging data (daily feeding, diaper changes, sleep, and so on) in a paper-based diary as the pediatrician suggested, they decide to install a high-end baby monitor on a crib, allowing them to remotely view high-quality video 24/7 (see Figure 3). As it is equipped with an infrared camera, the baby monitor captures Lucas' sleep and wake episodes at night. It also shows room temperature and the baby's day-to-day activities such as turning over, crying, and napping.

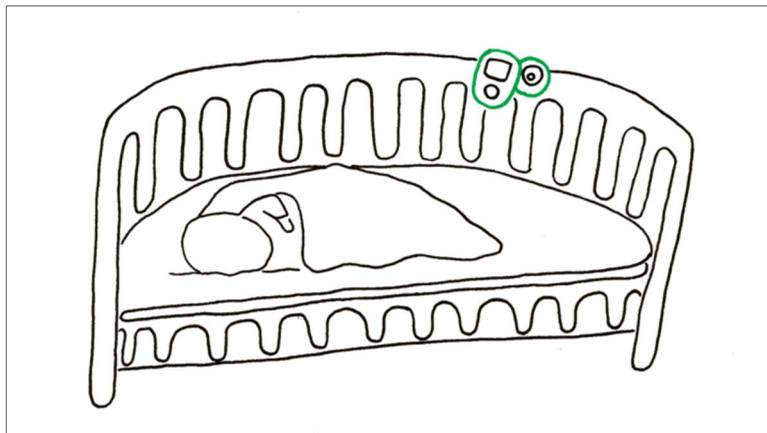


Figure 3. Scenario 3—Jessica and John installed a high-end video camera to capture baby's developmental data, which they hope to share with the pediatrician.

At their well-child visit, during which Jessica and John are encouraged to ask questions about Lucas, they are eager to share these additional data with their pediatrician. One of Jessica and John's main challenges, however, is to share the data and ask questions without overwhelming the pediatrician. To reduce the information overload while increasing the utility of the video data, they want data-capturing tools that help them create a succinct summary from the video. In addition to extracting video clips that capture important developmental activities, they need to be able to annotate the video summary using in situ tagging and ad hoc notes. Once the clinician sees the benefit of the additional video data, she can suggest other markers Jessica and John should look for.

RESEARCH CHALLENGES AND OPPORTUNITIES

As shown in all three scenarios, sharing of PGD between clinicians and patients can affect their personal interactions, workflow, and communication dynamics. While the outlook is positive, we should also attend to research challenges and negative consequences that could emerge in leveraging PGD in medicine.

Supporting Clinicians' Goals through PGD

In all three scenarios, clinicians are key partners in patient-clinician collaboration. As such, it is critical to involve clinicians during the early phase of design discussions, and their needs and goals should be reflected in design requirements and solutions. We developed these scenarios with limited participation from the clinician side, so they do not properly reflect clinicians' situations and perspectives. We assumed that clinicians are willing and motivated facilitators in introducing tracking to patients and leveraging PGD, but they may not be so willing to adopt technology and to discuss PGD with their patients due to workload perceptions and not understanding the possible benefits. Furthermore, they could have low data and visualization literacy,

which is important for effective data-driven communications. To successfully leverage PGD in patient–clinician collaboration, more work is needed to better understand clinicians’ goals and desires by working with them.

Facilitating Actionable Insight Gaining from Multiple Data Sources

Collecting personal data from multiple sources has become easier and more prevalent. However, as highlighted in scenarios 2 and 3, these data are scattered across many devices, apps, and platforms,³ making it difficult to get a holistic and synthesized view of one’s health and well-being. Despite the emergence of systems taking integrative approaches to data collection, such as AWARE (www.awareframework.com) and OmniTrack,⁸ and visualization, such as Visualized Self⁹ and Exist (<https://exist.io>), further research is needed to find easier ways for lay individuals and clinicians to gain insights from multiple sources of personal data.

Making PGD actionable in the clinic is crucial when improving patient–clinician interaction around self-tracking data. As in the scenario with Oliver, we envision a new role in healthcare, *the informatician*, who can help find actionable insights in heaps of life-data. Informaticians will be capable of using computer power and machine learning to discover and visualize not only correlations but also causations in data. They would play a key role in making incidental, multi-modal life-data accessible and useful in the clinical encounter.

Cultivating Sustainable Data Collection

The quality of data plays an important role in data-driven communications. To ensure data quality, we need to better motivate patients to collect their data diligently. First, we can make data tracking meaningful to them. For example, to engage people in tracking and self-reflecting on their fitness data, the Go & Grow system maps individuals’ walking data to the amount of water their plant receives.¹⁰ In addition, the possibility of creating a beautiful and unique artifact to express oneself using personal data can draw people to self-tracking. For example, Dear Data (www.dear-data.com) demonstrates that personal data can be visualized in beautiful, creative, and compelling, albeit subjective, ways, encouraging people to diligently collect data.

Ensuring the Clinical Relevance of Collected Data

Clinicians and patients might have very different perceptions about the value of PGD.¹ As it becomes easier to collect large amounts of data that might not be clinically relevant, clinicians are at increasing risk of data overload, which is likely to discourage them to adopt potentially useful PGD. To make PGD useful for clinicians, we need to make it easy and not time-consuming to take action.¹¹ Given that clinicians can judge the relevance and importance of PGD, there should be a mechanism in place to involve clinicians throughout patients’ data-capture process. In this way, clinicians can guide patients to capture markers that are relevant to their disease, hence increasing the value of PGD to clinicians.

CONCLUSION

In health and medical sciences, PGD bring a wealth of research directions and opportunities. In the future, tracking and sharing PGD might change the practice of clinical consultations as we know it: patients and clinicians have a data-driven medical consultation, improving patient engagement and speeding up the diagnosis. We hope to continue discussing ways to leverage PGD for better care with more clinician engagement, and encourage other researchers to contribute to future conferences including PervasiveHealth.

REFERENCES

1. H. Zhu et al., “Sharing Patient-Generated Data in Clinical Practices: An Interview Study,” *AMIA Ann. Symp. Proc.*, vol. 2016, 2016, pp. 1303–1312.
2. C.-F. Chung et al., “Boundary Negotiating Artifacts in Personal Informatics: Patient-Provider Collaboration with Patient-Generated Data,” *Proc. 19th ACM Conf. Computer-Supported Cooperative Work and Social Computing (CSCW 16)*, 2016, pp. 770–786.
3. E.K. Choe et al., “Understanding Quantified-Selfers’ Practices in Collecting and Exploring Personal Data,” *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI 14)*, 2014, pp. 1143–1152.
4. C. Kelley, B. Lee, and L. Wilcox, “Self-Tracking for Mental Wellness: Understanding Expert Perspectives and Student Experiences,” *Proc. 2017 CHI Conf. Human Factors in Computing Systems (CHI 17)*, 2017, pp. 629–641.
5. W.J. Korotitsch and R.O. Nelson-Gray, “An Overview of Self-Monitoring Research in Assessment and Treatment,” *Psychological Assessment*, vol. 11, no. 4, 1999, pp. 415–425.
6. L. Mamykina et al., “MAHI: Investigation of Social Scaffolding for Reflective Thinking in Diabetes Management,” *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI 08)*, 2008, pp. 477–486.
7. M.B. Rosson and J.M. Carroll, *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*, Morgan Kaufmann, 2001.
8. Y.-H. Kim et al., “OmniTrack: A Flexible Self-Tracking Approach Leveraging Semi-Automated Tracking,” *Proc. ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, 2017; doi.org/10.1145/3130930.
9. E.K. Choe et al., “Understanding Self-Reflection: How People Reflect on Personal Data through Visual Data Exploration,” *Proc. 11th EAI Int’l Conf. Pervasive Computing Technologies for Healthcare (PervasiveHealth 17)*, 2017, pp. 173–182.
10. F. Botros et al., “Go and Grow: Mapping Personal Data to a Living Plant,” *Proc. Int’l Working Conf. Advanced Visual Interfaces (AVI 16)*, 2016, pp. 112–119.
11. T.O. Andersen et al., “Aligning Concerns in Telecare: Three Concepts to Guide the Design of Patient-Centred E-Health,” *J. Computer Supported Cooperative Work*, pending publication.

ABOUT THE AUTHORS

Eun Kyong Choe is an assistant professor in the College of Information Studies at the University of Maryland, College Park. Contact her at choe@umd.edu.

Bongshin Lee is a senior researcher at Microsoft Research. Contact her at bongshin@microsoft.com.

Tariq Osman Andersen is an assistant professor at the Department of Computer Science at the University of Copenhagen. Contact him at tariq@di.ku.dk.

Lauren Wilcox is an assistant professor in the School of Interactive Computing at Georgia Institute of Technology. Contact her at wilcox@gatech.edu.

Geraldine Fitzpatrick is Professor of Technology Design and Assessment at TU Wien (Vienna University of Technology). Contact her at geraldine.fitzpatrick@tuwien.ac.at.