“Data Is One Thing, But I Want To Know The Story Behind”: Designing For Self-Tracking and Remote Patient Monitoring In The Context Of Multiple Sclerosis Care

Graham Dove
New York University
New York, USA
gd64@nyu.edu

Marina Roos Guthmann
New York University
New York, USA
mg5406@nyu.edu

Leigh Charvet
New York University
New York, USA
leigh.charvet@nyulangone.org

Oded Nov
New York University
New York, USA
onov@nyu.edu

Giuseppina Pilloni
New York University
New York, USA
giuseppina.pilloni@nyulangone.org

ABSTRACT

We report on design-focused inquiry into future multiple sclerosis (MS) healthcare; including a multi-stage design process with experienced MS clinicians, and formative evaluations with people living with MS. MS is a chronic, progressive, and unpredictable inflammatory neurological disease of the central nervous system that affects at least 2.8 million people worldwide. Walking impairments affect up to 85% of people diagnosed with MS. Responding to this, our focus is on design for longitudinally monitoring mobility, and in particular using wearable sensors that generate data on gait metrics to support clinical and self-care decision-making. We contribute to HCI research in three ways: (1) a detailed case study design process, including artifacts; (2) metaphorical framing concepts, with associated use cases illustrated through design scenarios; and (3) understanding of virtual-first practices in rehabilitation medicine that can be translated beyond MS care.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; **Empirical studies in interaction design**.

KEYWORDS

co-design, self-tracking, multiple sclerosis, collaborative healthcare

ACM Reference Format:
Graham Dove, Marina Roos Guthmann, Leigh Charvet, Oded Nov, and Giuseppina Pilloni. 2024. “Data Is One Thing, But I Want To Know The Story Behind”: Designing For Self-Tracking and Remote Patient Monitoring In The Context Of Multiple Sclerosis Care. In *Designing Interactive Systems Conference (DIS ’24), July 01–05, 2024, IT University of Copenhagen, Denmark*. ACM, New York, NY, USA, 22 pages. https://doi.org/10.1145/3643834.3661499

1 INTRODUCTION

In this paper we present a multi-stage inquiry into digital support for Multiple Sclerosis (MS) healthcare. We report on four rounds of design activities with experienced neuroligic physical therapists (PTs), as well as survey and formative design evaluations with people living with MS (PLWMS). Our work is framed by an overarching concern to support effective and timely decision-making through augmenting and contextualizing data about gait metrics generated during everyday activities. Similar to [70], we approach this through a lens of designing for future healthcare work. This perspective complements recent HCI research focusing on the self-care needs of PLWMS, and those affected by similar chronic neurological conditions [7, 8, 53].

1.1 Background to Using Wearable Sensors in Multiple Sclerosis Gait Assessment

MS is the most common inflammatory neurological disease of the central nervous system (CNS) among young adults [26, 73]. Affecting at least 2.8 million people worldwide [81], MS has an estimated prevalence rate in the US of more than 900,000 [80]. It is a chronic, progressive, and unpredictable condition that carries a great deal of uncertainty for PLWMS [1, 20], and also for clinicians who provide treatment [14, 38, 55]. Because of this, there is a strong desire for more and better information that can support increased understanding of the condition [44], effective personalized treatment and rehabilitation [36, 64], and a greater sense of agency [6]. Walking impairments are among the most frequent and debilitating symptoms, affecting up to 85% of PLWMS [40]. Yet, walking impairments can be responsive to treatment when identified early [15]. As a result, assessing changes in a patient’s gait has become a critical metric in routine clinical evaluations for MS. However, assessment and characterization of gait impairments typically relies on clinical observation of patients completing standardized tests, such as a timed 25-foot walk, measured manually with a handheld stopwatch during scheduled outpatient appointments [45, 66]. This can mean that clinical assessments lack timeliness and fail to accurately reflect community ambulation [46], while reliance on clinical appointments risks increasing disparities in access to care [28, 42].

DIS ’24, July 01–05, 2024, IT University of Copenhagen, Denmark

ACM ISBN 979-8-4007-0583-0/24/07
https://doi.org/10.1145/3643834.3661499

Quantitative gait analysis can help address limitations in standard mobility tests, but has typically involved specialized laboratories and complex equipment [34], resulting in restricted accessibility [11]. Wearable sensors with IMUs can help to reduce these requirements and augment observational assessments in clinics [27, 56]. They have shown repeatability and accuracy in estimating gait parameters like velocity and step length [59, 84], and when combined with machine learning (ML) have been used to identify motor progression [74]. However, in-clinic assessments (whether observational or based on data from wearables) offer a lens on understanding a person's gait in a standardized and controlled environment that is unrepresentative of functional mobility in community settings, at best providing a measure of capacity, rather than capability or better still performance [31]. Wearables also have potential to help bridge this gap between controlled clinical settings and quasi-continuous remote monitoring [11, 63], and thereby support more ecologically representative community gait assessments by helping identify changes in gait patterns and track the efficacy of pharmacological or rehabilitative interventions [57, 60, 79]. However, outside of clinical settings research has typically focused on measuring mobility capacity at a physical activity level [12, 84]. Moreover, the data wearables provide lack key contextual information necessary for explanatory power [70]. This is because mobility is complex, and among other things, is also reflective of a person's affective state and cognitive functioning, and responsive to interactions with the physical, social, and ambient environment [16, 77]. Increased contextual insight may result in more responsive interventions, and support personalized rehabilitation treatment [36, 64]; while better understanding the underlying cause of a particular symptom and its relationship (if any) to MS, can help support agentic self-care practices [6, 53]. In this paper we adopt a design perspective to investigate challenges and opportunities in using quantitative data about gait metrics, generated using wearable sensors, to provide insight into mobility and community ambulation in the context of MS care.

1.2 CONTRIBUTIONS

This paper contributes to HCI research in three ways. First, we present in detail a case study design process that includes remote collaborations with healthcare professionals and PLWMS, including documentation of our process and artifacts. Second, we present a set of metaphorical framing concepts to guide interaction design in this and similar contexts, and provide examples of their use in storyboard design scenarios illustrating how digital technologies might help monitor gait and mobility to better understand the progression of chronic conditions. Third, we contribute to HCI understanding of virtual-first practices in rehabilitation medicine that can be translated for use-cases beyond MS care.

1.3 Research Context

This research was conducted within the context of a collaboration between our university's schools of engineering and medicine, investigating changes to healthcare work introduced by digital technologies associated with telehealth and other virtual-first practices. It is situated in our university medical center, an urban academic institution and tertiary care centre, encompassing five hospitals and over 350 neighborhood clinics, and based in a large and densely populated city in the northeastern US. The authors of this paper are long-term collaborators experienced in HCI and biomedical research and design, as well as clinicians that treat patients with a range of neurological conditions including MS. None of the authors are living with a diagnosis of MS.

2 RELATED WORK

Similar to the work we present here, Seals et al. [70] investigates how quantified data about gait metrics might support the work of clinicians treating patients with MS. In contrast to this paper, Seals et al. approaches this challenge as primarily a data visualization task, creating a design prototype to probe clinicians' needs. Seals et al. finds that clinicians view these data as a potentially powerful lens into a patient's experience beyond the clinic, but would need additional context to support sense-making with regard to the data being visualized. For example, showing the activities being represented and related environmental conditions. Seals et al. also finds that clinicians' sense-making is likely to focus on a patient's functional performance, such as maintaining community ambulation and safety, particularly with regard to increased fall risk. Moreover, that this should be considered in the light of specific functional goals. Finally, Seals et al. highlights an important tension between visualizing data so that key information is immediately available to clinicians, while also creating opportunities for exploratory data analysis. However, limited consideration is given to the role of PLWMS beyond gathering and sharing gait and mobility data.

In contrast, Ayobi, together with various colleagues, presents a longitudinal inquiry into monitoring MS symptoms and experiences from the perspective of people living with the condition [6–8]. In an ethnographic study of existing practices that included disease monitoring, fitness tracking, and life-journaling, Ayobi et al. [8] highlights how important it can be for PLWMS to try to gain a sense of control over the disease. This was reported as a strong motivation for starting self-tracking practices, which could lead to positive outcomes with regard to self-care. Building on this initial study, Ayobi, Marshall and Cox [7] highlights the role that personally meaningful metrics and representations can play in helping PLWMS build a greater sense of agency over the disease; and suggests that customizable and pictorial self-tracking tools can promote mindfulness in ways that predefined tools focusing on normative indicators may not. Ayobi [6, p.145-148] further suggests that designing for individual agentive capacities is important as it respects peoples' changing priorities and relationships to self-tracking activities. It should be noted that this body of work focuses on self-motivated people with a preexisting interest in self-tracking practices. Also that the relationships of participants with their healthcare providers receives limited consideration.

Wendrich et al [83] also presents an interview study exploring experiences and perceptions of self-tracking with PLWMS. In this study, which pays closer attention to relationships with healthcare providers, participants again suggest a need for self-tracking tools that can be adapted to personal situations, but also indicate a desire that self-monitoring data be fully integrated into treatment plans. This, the authors suggest, would result in better informed clinicians who could offer guidance on how data generated through self-tracking might provide value to patients. Building on this initial work, Wendrich and Krabbenborg [82] presents an interview study with participants in a longitudinal MS self-tracking program. This study finds that participation in research, and a desire to contribute knowledge that will benefit other people living with MS, could be stronger motivations for adherence to the program than personal self-management of MS symptoms. Moreover, that this motivation remains persistent despite participants noting the inconvenience of self-monitoring tasks, and the emotional burden of self-monitoring being a reminder of MS.

Marziniak et al. [51] reviews literature on clinical trials of remote MS healthcare, including symptom tracking, with a focus on the efficacy of the interventions that were trialed. The authors note that the high-level of technology adoption by PLWMS in the US and Europe offers opportunities for improved care. However, similar to Wendrich and Kabbenborg [82], they also suggest that for PLWMS repeated self-tracking tasks can can feel intrusive and act as an unwelcome reminder of the condition. Marziniak et al. also suggests that the need to manage and respond to high volumes of data generated by self-tracking practices may discourage widespread clinical adoption.

Nunes, Andersen and Fitzpatrick [53] investigates in-person and online interactions that people living with Parkinson's disease (a chronic neurological condition manifesting symptoms that share similarities with those observed in MS) have with their clinicians. The authors contrast findings that highlight the agentic role people living with Parkinson's play in their own healthcare with a review of contemporary self-tracking technologies, and highlight opportunities for richer participation in shared decision-making. In prior work, Nunes et al. [54] presents a review of HCI research into self-care technologies used by people living with a range of chronic conditions. Self-care in this research is broadly defined as being concerned with managing symptoms, treatment, emotions, and lifestyle changes, influencing the course of a condition through practical everyday actions, and maintaining a satisfactory quality of life for as long as possible. The authors highlight a series of design tensions, in particular those between *patient autonomy* and *clinician control*, and devices focused on *medical symptoms* or *lived experience*.

In the following sections we build on this prior literature by presenting our inquiry into how wearable sensors might support greater understanding of how MS impacts mobility, leading to walking impairment. We return to possible tensions between monitoring and self-tracking, from the perspectives of providers and patients, in our discussion section 7.

3 METHOD

We present design inquiry into digital support for MS healthcare, focusing on the use of wearable sensors to generate data on gait metrics and provide insight into walking impairment; thereby supporting clinical decision-making and agentic self-care. The process we describe was guided by prior literature on research through design (RtD) [86, 88] and co-design [67], which describes approaches that adopt design practice methods and capture knowledge in the artifacts created. To report on these processes, we carefully describe key activities at each stage, outline our analysis and findings, and introduce metaphorical framing concepts (see section 3.1). These framing concepts were used to help bridge analytical and generative processes, and in guiding design choices by characterizing some of the potential roles for digital technologies in MS healthcare. They help to support key activities throughout, in particular we used them in helping to document and reflect on design judgements [10, 24], and to identify preferred states and make explicit new framings as they emerge [21, 89].

In section 4, we report on a multi-stage design inquiry process undertaken with clinicians from our university medical center. This includes two rounds of interviews (see sections 4.1 and 4.2), a generative co-creation activity using a set of digital healthcare design cards we created for the project (see section 4.3), and an early-stage design evaluation using a speed dating approach with storyboards representing three design ideas (see section 4.4). Each of the methods, including our approach to analyzing the data collected, is described in detail as part of our reporting on the particular activity in question. Following this, in section 5, we report on initial work investigating the perspectives of PLWMS. This includes analysis of data from public Reddit forums, a survey of PLWMS, and interviews and formative evaluation sessions conducted with PLWMS. Again, each of these methods, including the analysis approach, is described in detail when we report on the activity in question.

The starting point for our inquiry was an interest in future practices in MS healthcare. It has been framed by our prior research interviewing diverse neurological clinicians, and findings reported in HCI literature. In particular, Seals et al. [70] highlight: (1) the importance of fall risk to patient safety and community ambulation to patient independence; (2) the need to contextualize gait data with a patient's individual circumstances, everyday activities, and particular treatments; and (3) opportunities to support collaborative goal-agreement and decision-making. It is additionally informed by HCI research indicating that self-tracking can support self-care practices and help build a sense of agency and control [6, 53]. This research was conducted under the approval of our university's Institutional Review Board (IRB-FY2021-5295).

3.1 Framing Concepts

At each stage of our design inquiry we generated metaphorical framing concepts as a way to represent findings and help guide design judgements. These are first presented and explained within each design activity (e.g., 4.1.2). A full listing is then provided towards the end of this paper in Table 1, as part of our Summary and Synthesis (section 6) where they are illustrated in storyboard scenarios; and where we show in more detail how these framing concepts are represented in the design ideas that were selected for storyboarding. These framing concepts allow us to characterize different roles that technologies might play in supporting the work practices of clinicians and the self-tracking practices of PLWMS. We think of these framing concepts similarly to the way that Lawson and Loke [43] used framing concepts to describe different roles that an artificial intelligence (AI) might play in computer aided design. Because they are intentionally ambiguous and low on prescriptive detail, these framing concepts allow us to explore and evaluate different options, and suggest 'how might we?' questions during ideation.

Depending on the situation, we can respond to these concepts in ways that are either more speculative or more pragmatic. Our use of these framing concepts builds on prior research showing: (1) that metaphors help designers to translate between familiar design types and particular situations [61, 68], and perceive new features relevant to the problem at hand [69]; (2) that shows metaphors can act as a representational tool, to show how a complex artifact is used, and as a rhetorical device that invites deeper thinking or reframing [23]; and (3) that highlights how during co-design activities metaphors can provide common ground where domain experts and designers meet and share ideas [25].

4 DESIGN WITH PHYSICAL THERAPISTS

In this section we report on a design inquiry in which we collaborated with three experienced physical therapists (PTs) recruited from different practices across our university's medical center network. We recruited PTs because they have particular skills, knowledge, and experience in focused and hands-on rehabilitation treatment that extends the expertise within our research and design team. Each of the PTs are experienced in treating patients with MS and similar neurological conditions, but each works in a very different clinical context. Over a period of nine months, we met with the PTs on four separate occasions, twice in a group context and twice individually. Each of the sessions took place remotely over Zoom. Each session lasted one hour and meetings were recorded. An initial transcript was created automatically, and subsequently cleaned and revised. During this nine month period, our design team also met online via Zoom on a weekly basis and occasionally in-person for additional design activities.

- *Participant PT1* is a specialist in neurologic physical therapy with more that 15 years experience. PT1 works in a hospital setting where people are receiving in-patient care. In the case of people living with MS this is typically because they are experiencing a symptom flare or have been admitted because of a comorbidity.
- *Participant PT2* is a specialist in neurologic physical therapy and practice manager with more than 15 years experience. PT2 works in an adult outpatient physical therapy clinic where patients are typically receiving ongoing rehabilitation for neurological conditions such as MS, or following a stroke or spinal chord injury. PT2 treats people living with MS who have a variety of MS associated symptoms at different stages of severity.
- *Participant PT3* is a specialist in orthopedic physical therapy and practice supervisor with more than 15 years of experience. PT3 works in an outpatient clinic where the focus is on people who have an active lifestyle, and are likely to visit due to sports injuries or similar problems. In this context, MS may be a confounding or complicating issue, or the undiagnosed cause of underlying problems.

4.1 Activity 1: Understanding Current Practice

The first activity was a group interview and discussion in which the first author probed on the protocols that the PTs follow in their different practices and with different patients. Having introduced themselves and provided a brief professional background, each PT was first asked to describe the type and diversity of patients they treat, with a focus on people living with MS and similar neurological conditions. For example, whether the patients were likely to be newly diagnosed, how severely their symptoms present, their mobility and degree of independence, etc. Following this, each PT was asked to describe in detail what an initial appointment with a new patient might look like. For example, descriptions of the location and setting of the appointment, how the PTs prepared, the type of assessments undertaken with the patient, the metrics recorded for these assessments, the process of agreeing objectives and goals with patients, and the patient-clinician communications that follow. We repeated this process discussing PTs' recurring sessions with a patient. This initial questioning took around 40 minutes. For the remaining 20 minutes the first author facilitated a group discussion in which we attempted to unpick potential similarities and differences between practices and the challenges PTs face.

4.1.1 Analysis. We re-watched the recording multiple times, amending the automated transcript as necessary. We then created an affinity diagram [32] to identify, map, and visualize, key insights and their relationships. This identified four key findings that informed our first framing concept.

4.1.2 Findings: (1) PTs will typically focus on functional assessments and a patient's ability to complete everyday tasks. This influences the goals and exercise programs they agree with patients. (2) PT clinics have limited space, and there will likely be multiple assessments taking place at the same time. (3) PTs need to offer support and guarding to patients, and to be present and aware of the risk that a patient might fall during an assessment exercise; so avoid asking them to take notes or use handheld devices to record analog measurements. (4) Results from PT assessments may not be available until after the initial appointment is over, and so patients may need an additional consultation.

In this session, PT2 drew attention to the constraints and complications of their clinical setting by first saying, *"A confounding variable to that is we have a very busy clinic with 17 therapists working. Okay, and one walkway and sometimes the patients are sharing a walkway and you're trying to step out the way. You're not supposed to give the person encouragement too much. You're not supposed to really tell them too much, except for times. There's very strict instructions for how this test is performed you tell them that, where they are in the six minutes at the end of each minute. So you know, we have to count them and sometimes with a lot going that actually gets more difficult than it seems."* And following this up later by saying, *"I think barriers for us are like literally timing and guarding. Doing some of those things are real barriers right, so if someone needs a little bit of close assistance we're trying to hit the stopwatch. You know we're wearing it around our necks so our hands can be free, but then you're also getting to it at the same time, while guarding"*. In response to this activity, we were particularly prompted to consider the physical hands-on requirements of working as a PT, and whether the wearable sensors and digital technologies we have been considering as support for longitudinal and remote insight might also be used to support hands-free measurement that could provide insight during in-clinic appointments. This potential reframing opened up new areas for consideration.

Figure 1: The 4 example wearables that we presented to PTs during Activity 2, these were: (A) a wearable used by runners that is attached to shoelaces; (B) a pressure-sensitive system for innersoles; (C) a clinical sensor for balance and gait worn on a belt around the waist; and (D) a consumer smartwatch.

*4.1.3 Framing concept: **'The ideal assistant'** helps the clinician maintain a clear and present focus on their patient, and supports them with timely information.* This concept reflects that PTs' actions and attention are primarily focused on providing physical care and support for the patient, that taking assessment metrics can be challenging, and that detailed reports of assessment data may be delayed.

4.2 Activity 2: Exploring Technology-Supported Future Practice

Our second activity was again a group interview and discussion facilitated by the first author. This time we focused on how digital technologies might help record assessment metrics and capture longitudinal data between a patient's visits to the clinic. We first identified specific metrics that would be valuable to PTs and discussed current challenges to recording them. Example metrics included: time-taken to complete an exercise, which is currently measured with a stop watch; distance-travelled within a set time, which is currently measured with floor markings or a tape; gait speed and balance, which are currently observational measures on a qualitative scale; and fatigue and pain, which are currently self-reported by patients using qualitative scales. Constructs such as risk of fall or community ambulation, which are judgements derived from PT assessments and conversations with the patient, were also discussed.

We then showed PTs four examples of commercially available wearables that might be used in self-tracking: (1) sensors worn on shoelaces; (2) sensors attached to a shoe innersole; (3) a sensor worn around the waist; and (4) a smartwatch worn on the wrist (see Fig. 1). This was followed by a group discussion that focused on the potential benefits and challenges of each sensor, and in particular potential implications of where and how the sensor is worn. For example, a sensor worn around the waist is close to to the persons center of gravity, while a sensor worn on shoelaces might be challenging to fit, and a sensor on an innersole might need to be added to multiple pairs of shoes. Finally, we asked each PT to select one of the wearables and describe how they might use it with a patient.

4.2.1 Analysis. Similarly to activity 1, we first re-watched the video recording multiple times and amended the automated transcript; and then created an affinity diagram to highlight key insights. Following this, we conducted an inductive thematic analysis [13] of the combined transcripts from activity 1 and activity 2, in order to identify cross-cutting patterns and themes. This analysis was carried out individually by the first and second author, who then came together for agreement reaching, with the last author providing an additional assessment of the derived themes for confidence.

4.2.2 Findings. (1) Tools that support small tasks and help in preparing for daily activities, or which lead to incremental improvements such as optimizing processes and documentation, can make an important and impactful difference to the clinical experience PTs offer. (2) In-clinic measurements are artificial, because the behavior of others (e.g., clinicians, staff, visitors, etc.) is different in this environment (i.e., they get out of the way), and floors are typically flat and without curbs or similar obstacles. In addition, there are no variations in weather conditions (e.g., heat can impact PLWMS, and wet or icy conditions increase the risk of falling). This means that evidence to extrapolate performance into functional ability, or for use in selecting interventions and exercise plans, is limited. (3) Data from wearables offer fine-grained measures that could help identify compensations the patient might be making (e.g., when walking) or in predicting an increased risk of falling. However, this level of detail is in tension with PTs' current preference for functional improvements that focus the patient on quickly returning to important everyday activities. (4) Longitudinal data can provide motivation for or validation of a patient's functional improvement, and so influence clinician and patient attitudes and inform education; for example, investigating the connection or mismatch between self-reported levels of fatigue and patterns in activity level recorded by wearables.

During this activity PT3 highlighted the difference between in-clinic measurements and the potential for data collected using wearables in everyday life, saying: *"It's a truer representation of what they're experiencing in life, you know, because even the things we measure in clinic are somewhat artificial. Everybody kind of knows, to give other patients right away, you know our floors are completely flat, and you know it's good in the sense that it keeps you safe but*

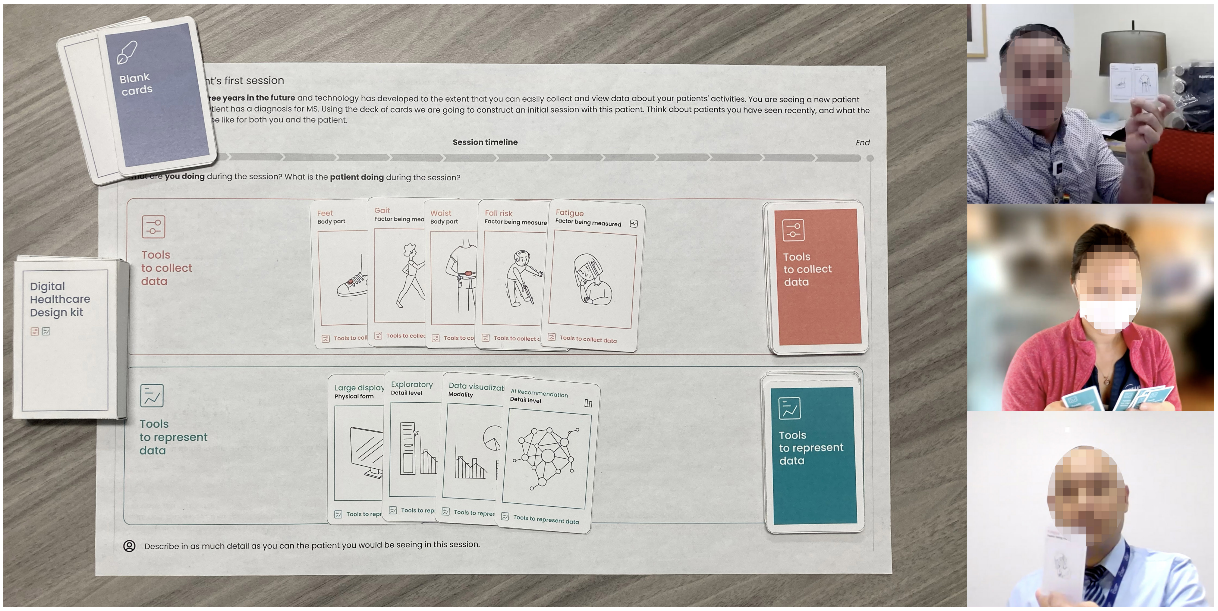


Figure 2: A set of the digital healthcare design cards used during online co-design activities with physical therapists, with insets showing therapists selecting cards and holding them up for the camera.

it's a little bit artificial it doesn't mimic the uneven sidewalks in the city and the curbs that they have to step down and so on, the black ice that they have to avoid". In response to this activity, we were particularly prompted to consider tensions between the ideal of fine-grained nuanced understanding of a patient's mobility and support for PTs' work improving functional capacity. In chronic debilitating conditions such as MS this tension may be heightened, because maintaining functional abilities or slowing their loss can in themselves be major achievements. PTs use data from standardized assessments to gain insight into functional aspects of mobility, such as risk of fall and community ambulation. Currently this is limited by infrequent in-clinic measures, but it offers a template for thinking about how tensions might be eased.

*4.2.3 Framing concept: **'The data interpreter'** uses insights from outside the clinic to monitor change and help resolve tensions between short-term aims and longer-term consequences.* This concept reflects that PTs would love to better understand what happens outside the clinic between visits, as this is more important yet may be very different from what they see during appointments.

4.3 Activity 3: Co-Creation Using Digital Healthcare Design Cards

Design is a process of making sense of things [39], and of devising courses of action aimed at changing existing situations into preferred ones [72]. In co-design processes this typically includes designers facilitating and guiding participants through a process of joint imagination in order to uncover opportunities that are not readily expressed through interview and conversation [67, 75]. To explore possible futures more speculatively, we worked with each PT individually, aiming to explore positive visions and to put minor challenges to one side. This third activity followed a co-creation protocol (see section 4.3.2), using a custom card-based toolkit that we created for the project, and which consists of 38 prompt cards and two worksheets (see Fig. 2 and section 4.3.1 for more detailed descriptions of the toolkit, and Appendix 1 for a listing of cards). Prior to each session, a kit was printed and constructed, and sent to the PT. The second author was the primary facilitator in these sessions, with the first author probing with follow-up questions where appropriate.

4.3.1 Digital Healthcare Design Cards. The toolkit we designed consists of thirty-eight cards and two scenario worksheets. It was informed by findings from activities 1 & 2, and our prior experiences researching and designing digital healthcare support. It builds on prior use of card-based methods as ideation and creativity prompts e.g., [29, 47]. There are two main categories of cards, focusing on how to *collect data* and how to *represent data*. The cards depicting data collection are sub-categorized into three groups: (1) the factor being measured; (2) the method of measurement; and (3) the part of the body in focus. The cards depicting data representation are sub-categorized into: (1) the modality of representation; (2) the form factor of the device used for representation; and (3) the detail level at which data will be presented. We also included four blank cards in each category, to represent PT suggestions. The kit also includes two scenario worksheets to help structure ideation: (1) for an initial in-clinic appointment with a new patient; and (2) for longitudinal relationships with recurring remote interactions.

4.3.2 Co-creation protocol. We first introduced the toolkit, explaining each of the cards and offering PTs an opportunity to ask questions. We then introduced the design scenario, and asked PTs to imagine they were three years into the future and everyday irritations that cause technologies to not perform as expected have been ironed out. We selected a positive near future scenario in

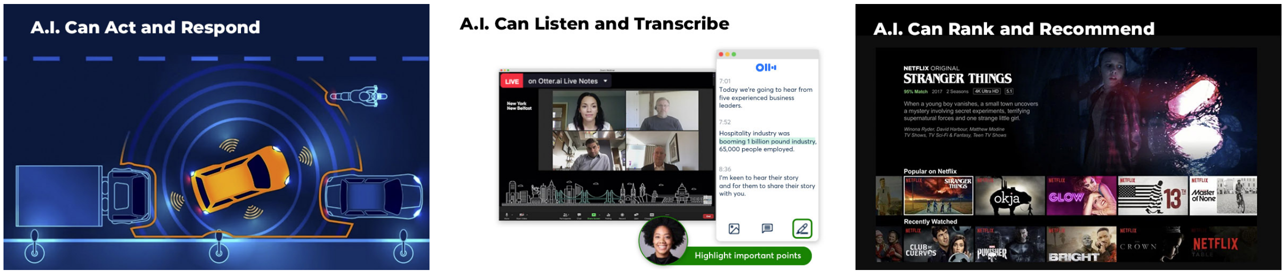


Figure 3: Examples of AI capability prompts used during ideation for design scenarios. From left to right these remind us that AI can be used to respond to sensors, to capture and analyze speech interactions, and to make recommendations based on patterns of prior information.

order to explore existing, and potentially familiar, technologies in a context where commonly experienced challenges are put to one side. PTs were then asked to imagine an initial clinical appointment with a patient they are seeing for the first time. We asked them to describe how an ideal experience might be created, using the cards they selected as a prompt for discussing different technologies. We encouraged PTs to select multiple cards and explore different alternatives at each stage. We also prompted them to use the blank cards to introduce technologies we had not included; which each PT did on multiple occasions. PTs were asked to hold up each card they selected and show it to the camera, and to explain the role of their selection in the 'ideal' appointment. We then probed on the pros and cons of their selection. After PTs completed the first scenario, we probed more deeply on their selections to clarify how their ideas should be interpreted. We then introduced a second scenario, and repeated the co-creation process, this time focusing on recurring interactions, including remote data collection and virtual communications, as well as in-person activities. We prompted PTs to consider things that patients and clinicians might currently have difficulty sharing or communicating, and to think about how we might create ideal longitudinal relationships. PTs were asked to select a variety of cards and use blank cards to introduce alternatives, and again this was followed by deeper probing on how ideas should be interpreted.

4.3.3 Analysis . We re-watched the video recordings of each co-creation session multiple times, updating and amending the automated transcript, and taking individual notes of things we considered important; such as selecting sensitizing quotes to reflect PTs' priorities. The design team then held an in-person ideation session to synthesize findings. During this session we referred to notes and transcripts, revisited the video recordings, and discussed our selected quotes. We also wanted to stay sensitized to the opportunities that artificial intelligence (AI) offers towards mitigating increased workloads that digitally-supported remote and longitudinal care can make on on clinicians [9, 50], and so we created a series of AI capabilities posters similar to those discussed in [85], which were pinned up and visible as prompts. For example, "AI can listen and transcribe" showing Zoom, "AI can rank and recommend" showing Netflix, and "AI can estimate and predict" showing the Redfin real estate website (see Fig. 3). Our analysis identified seven findings, and two framing concepts. We iterated and refined our ideas, settling on three design scenarios (see section 4.3.6).

4.3.4 Findings . (1) Conversation, and hearing first hand what the patient experiences in everyday activities, like going to the grocery store, helps PTs understand their abilities and needs, and set goals. (2) PT goals are typically based on functional capabilities, and set with reference to standardized test outcomes (e.g. balance and gait) and qualitative self-reporting (e.g. fatigue). (3) Best practice prompts or reminders, and automated comparisons to normative data, can guide PTs' decision-making. Insights from data can also be linked to motivating messages and adherence reminders for patients. (4) While data in community settings offers new insight and opportunities to support self-management, not all data are equal. For example, ninety-nine times out of one hundred someone may experience no difficulty getting up out of a chair, but if that one time happens in public in a restaurant it can have a significantly out-sized impact. (5) MS is a chronic degenerative condition in which capabilities differ between people and over time, and so the value and usability of different wearables is contingent on the particular situation. (6) Tensions can persist between the way a patient feels and indicators found in data, such as when assessment scores remain consistent but a patient perceives an improvement or deterioration. (7) Data and communication should be retained in the medical center's electronic health records, because other systems fall outside regulatory requirements, and limited time is available to view data.

4.3.5 Framing concepts: **(1)** ***'The facilitator'*** *helps to open up patient-clinician conversations, scaffold decision-making, and prompt considerations that might otherwise be overlooked.* This concept reflects the importance of conversation to PTs, and the role that insight from data might play in supporting collaborative decision-making. **(2)** ***'The motivator'*** *informs and encourages PLWMS, and supports adaptive responses to changing circumstances.* This concept reflects the tension between time-limited PT care and the experience of living with a chronic condition.

4.3.6 Design scenarios. Our ideation session resulted in three design scenarios that each consider different temporal points in the patient-clinician relationship, different activities, and different support technologies. They illustrate speculative implementations of the different framing concepts generated through activities with PTs. While it had not initially been our intention to create an overarching flow to connect each scenario, it became clear during refinement that they could be temporally ordered. Reflecting on this later, we found it was important as it helped PTs draw connections and make comparisons across scenarios.

APTA (Assisting Physical Therapy Autonomously): A scenario in which wearable sensors, cameras, and AI, are used to support physical therapists during in-clinic sessions. The aim in this scenario was to explore how technology might help in automating data collection and speed up analysis, offer recommendations and support planning, and reduce preparatory and post-session overhead. This scenario incorporates *The ideal assistant* and *The facilitator* concepts.

TAMPA (Training and Motivating Patients Autonomously): A scenario in which wearable sensors, smartphone cameras, and AI, are used to support PLWMS in completing and fine-tuning at-home exercise programs. The aim with this scenario was to explore how technology might offer support for interventions and provide improved insight into their effectiveness. This scenario incorporates *The motivator* and *The data interpreter* concepts.

Fall Risk Monitor: A scenario in which wearable sensors, self-reporting, and AI, are used by people living with MS and their clinicians to gain insight into mobility in everyday situations (see Fig. 4). The aim in this scenario was to explore how self-tracking technology might help support functional mobility and also offer fine-grained insight. This scenario incorporates *The data interpreter* and *The facilitator* concepts. Fig. 4 starts by illustrating how data might be collected in everyday settings, a key wish for PTs to help in their understanding. This includes quantitative data about gait metrics (e.g., blocks 3-6 and block 9) as well as brief experience sampling prompts to probe on other factors they were concerned about (e.g., block 7 and block 10 'End of day'). Switching to the PT and AI view shows how *The data interpreter* and *The facilitator* are combined within interactions between Billie and the AI, and Billie and Alex, and how the data and AI insights offer common ground for Billie and Alex during their virtual appointment (e.g., blocks 15-17).

4.4 Activity 4: Speed Dating With Storyboards

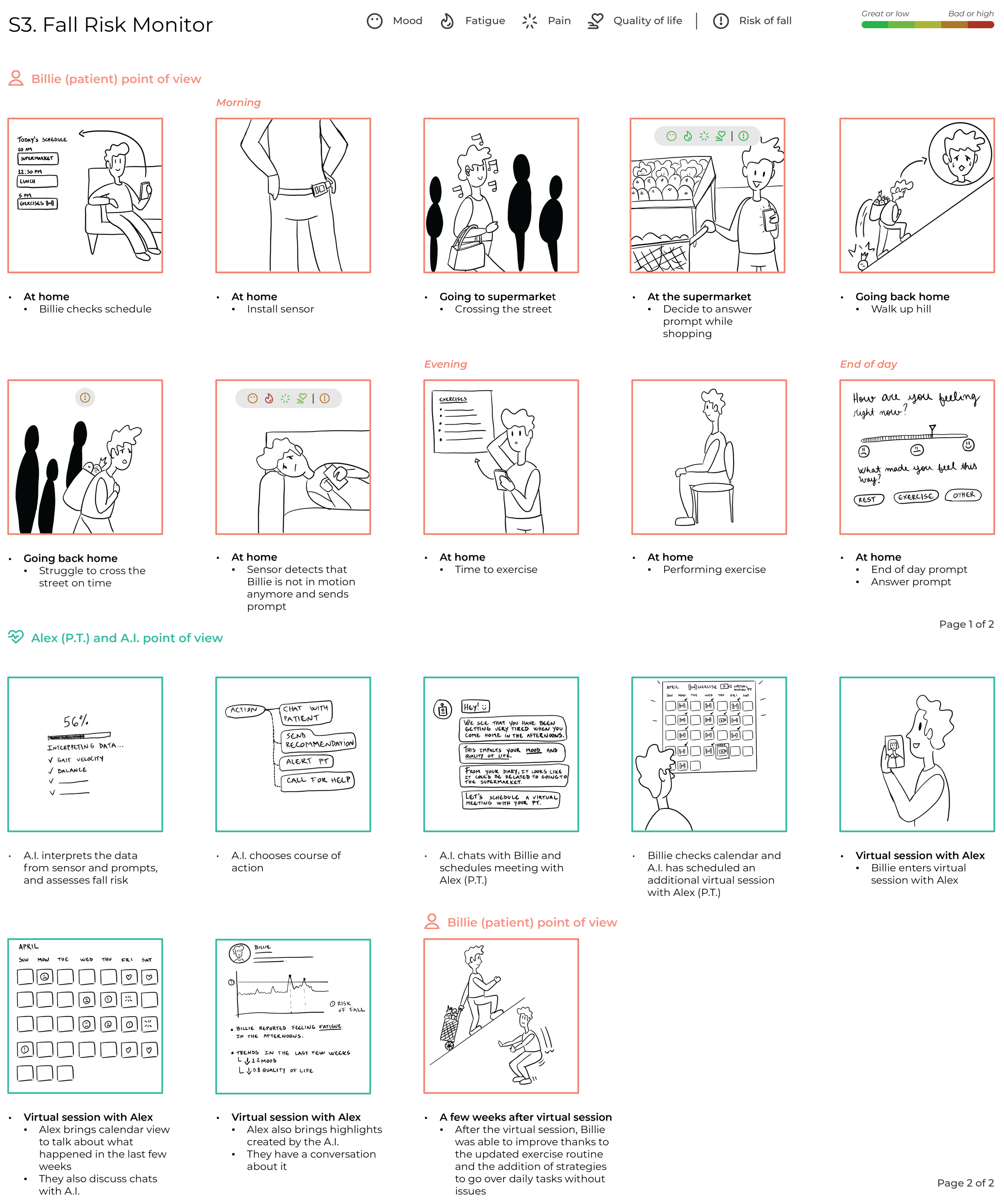
The risk that design ideas will struggle for adoption or acceptance is greater in situations where few design patterns exist, or when speculative futures are being considered. Speed dating design ideas with stakeholders, using low-fidelity representations, can reduce these risks and also reveal opportunities [19, 87]. For our fourth activity, again working with PTs individually, we created storyboards representing each of the design scenarios from section 4.3.6 (see Appendix 2 for additional storyboards). Prior to these sessions we printed copies of each storyboard onto 17" x 11" paper, which we sent to the PT in a package that also included small colored stickers. In the sessions, 15 to 20 minutes was spent on each scenario in turn. We asked PTs to read through the storyboard block by block out loud and place colored stickers to indicate something they particularly liked, something they didn't like, or something that raised a question. Following a think-aloud protocol, we prompted PTs to verbalize their thoughts and feelings, and explain why they were adding a sticker. After each storyboard walk-through we went back to probe more deeply on key points and outstanding questions. Once all three storyboards were completed, we had a final wrap-up to ask cross-cutting questions. Sessions were facilitated by the first author, with the second author supporting and probing where appropriate.

4.4.1 Analysis. We re-watched the video recordings multiple times, amending the automated transcript as necessary. We then conducted a debrief and reflection session in which we discussed our notes and re-visited key segments of the video from each session, paying close attention to each PT's use of colored stickers. This process helped us to understand where our scenario interpretations seemed validated and where we might be wide of the mark. We did this, with a focus on consolidating and moving forward in developing design ideas. Our analysis resulted in six new findings and one additional framing concept.

4.4.2 Findings. (1) AI should be used to save time and administrative overhead, or standardize best practices and compliance e.g., automating measures, transcribing conversations, and entering notes into the electronic health record (EHR); or suggesting adjustments to assessment plans, and quickly collating results with comparison to normative data. (2) AI should not reduce PTs' autonomy and agency, but rather support critical thinking in their role as the key decision-maker. (3) Mobility and balance data from wearables, in combination with self-reported fatigue and pain, and environmental information, can offer insight into how effective interventions are outside the clinic, which is where everything important to the patient happens. (4) AI insights could help connect the dots for patient-clinician conversations, and support planning for mitigation strategies; but only where it is trusted enough for PTs to offload additional data work. (5) PT patients will inevitably finish a program and leave, which can be a source of stress to them. AI motivation and adherence support, especially when connected to educational support, could help this transition. (6) In practice it is too challenging to set up cameras and video for feedback, which requires time that is probably not available.

This activity prompted us to to reflect on where PTs might find AI automation genuinely helpful. In the clinic this is where it can save time and administrative overhead, or where it could be used to help with best practices standardization and compliance or support critical thinking; and in particular where it can reduce the need to manually take and record measures, because they are attending physically and mentally to the patient while the assessments are being completed. However, greater value is likely found in AI generated insights from outside of the clinic, where 'everything that's important to the patient happens', as these might help in better understanding the impacts and effects of intervention programs and support collaborative planning of mitigation strategies. The PT's objective is to have patients finish their program and leave, which can be stressful for the patient. AI also offers opportunities to help in this transition, and support adherence to agreed plans e.g., through tailored educational support.

*4.4.3 Framing concept: **'The community-based assistant'** helps bridge the gap between home and clinic, manage day-to-day interactions, and provide reliable and easy to digest insights and advice.* This concep reflects how AI might be considered a partner in the care process, helping to interpret and act on data the PLWMS is generating using wearables in contexts that are important to them.



- **At home**
  - Billie checks schedule
- **At home**
  - Install sensor
- **Going to supermarket**
  - Crossing the street
- **At the supermarket**
  - Decide to answer prompt while shopping
- **Going back home**
  - Walk up hill
- **Going back home**
  - Struggle to cross the street on time
- **At home**
  - Sensor detects that Billie is not in motion anymore and sends prompt
- **At home**
  - Time to exercise
- **At home**
  - Performing exercise
- **At home**
  - End of day prompt
  - Answer prompt

- A.I. interprets the data from sensor and prompts, and assesses fall risk
- A.I. chooses course of action
- A.I. chats with Billie and schedules meeting with Alex (P.T.)
- Billie checks calendar and A.I. has scheduled an additional virtual session with Alex (P.T.)
- **Virtual session with Alex**
  - Billie enters virtual session with Alex
- **Virtual session with Alex**
  - Alex brings calendar view to talk about what happened in the last few weeks
  - They also discuss chats with A.I.
- **Virtual session with Alex**
  - Alex also brings highlights created by the A.I.
  - They have a conversation about it
- **A few weeks after virtual session**
  - After the virtual session, Billie was able to improve thanks to the updated exercise routine and the addition of strategies to go over daily tasks without issues

Figure 4: The 'Fall Risk Monitor' design scenario, an example storyboard used during speed dating activities with PTs. This illustrates *The data interpreter* and *The facilitator* concepts, reflecting how self-tracking might help PTs better understand what happens outside the clinic between visits and how insight from data might support collaborative decision-making.

5 LEARNING FROM PEOPLE LIVING WITH MS

Continuing care is a collaboration between clinician and patient, among others, and so we also wanted to learn from people living with MS (PLWMS) to build on our work with PTs. To gain a wider perspective, and avoid potential conflicts of interest, we chose to focus our research through the Reddit social media platform rather than recruit from patients at our university medical center. First we analyzed public Reddit posts on subjects related to self-tracking. We then conducted an online survey, recruiting through the subreddits we analyzed. Finally we conducted formative evaluation interviews, recruiting from participants who had completed the survey.

5.1 Public Reddit Forums

In recent years, the social media platform Reddit has become widely used as a source for research data, and as a place for recruitment to survey and interview studies. This is due largely to it's subreddits that support communities focused on particular topics of shared interest or concern [2, 62, 71]. To gain some initial insight into how PLWMS engage with wearable devices and self-tracking, we identified 4 subreddits where experiences are shared publicly: 'r/MultipleSclerosis', 'r/MultipleSclerosisWins', 'r/MultipleSclerosisLife', and 'r/MultipleSclerosis_MS'. Our selection took into account activity levels on the subreddit, including subscriber count, and frequency and volume of engagement (measured by posts and responses). This information was obtained from subredditstats.com.

5.1.1 Analysis. Using the Python PRAW (Python Reddit API Wrapper) library to connect to the Reddit API, we retrieved a list of thread topics, together with 'post id', 'author', 'title', 'self text', 'count of comments', 'upvote count', 'upvote ratio', and 'flair' fields. This generated a data set of 1,396 initial thread posts that covered the period November 2012 to May 2023 (NB. 'r/MultipleSclerosis' is by far the largest and longest standing of the subreddits and so provided the majority of posts). We conducted a simple key word search for the terms 'track', 'monitor', and 'quantify', from which we identified 8 threads that discussed self-tracking and monitoring symptoms. We then read through the full list of thread titles in order to check we had not missed anything obvious, looking more closely at the message text where it might be concerned with self-tracking. This identified one additional post on the topic of 'self-advocacy'. We then pulled all of the replies to these comments, a total of 168 comments. We first did a quick read through each of the comments to see which threads were relevant to our topic, which removed 3 of the smaller threads and left us with 125 comments (NB: the final data set was made up of comments that all came from 'r/MultipleSclerosis'). These we read closely to identify informative themes, resulting in six findings.

5.1.2 Findings. (1) Tracking MS symptoms was not a topic raised frequently within these forums. (2) Tracking MS symptoms was most often discussed as requiring a manual record, as each person's needs are different. (3) Wearables are used to track sleep and/or activity, or to monitor temperature, reflecting the impact that heat and fatigue can have on PLWMS. This was discussed with respect to adjusting behavior or being prepared for 'bad days'. (4) Reassurance from fall monitoring was an important motivation for having a wearable, but was discussed as preparation for the future rather than a current need. (5) Wearables and self-tracking data were also used to highlight achievements, and show that a diagnosis of MS does not become the defining factor about a person. (6) For some PLWMS who contributed to the subreddits, tracking symptoms was viewed as an unnecessary reminder of MS that offered no real benefits.

5.2 Formative Evaluation Survey and Interviews

We posted to each of the subreddits recruiting for survey participants. The survey asked for simple demographics and time since the respondent's MS diagnosis. It then probed on how the respondent experiences a range of MS symptoms and whether they currently track these or would like to track them in future. We also included a free-text field for respondents to provide any further details they wished, and asked whether they would be willing to evaluate and discuss our design ideas. We received a total of two hundred and twenty six completed surveys. However, during our analysis we noticed that two hundred and four were made within a single four hour period and contained repetitive similarities in the free text field. Putting these responses to one side as a precaution, we were left with twenty-two completed surveys that we read closely and compared to our subreddit analysis, providing two additional findings:

5.2.1 Findings. (1) Around one quarter of respondents currently tracked one or more MS symptom, and a similar number replied that tracking is not helpful. Half of the respondents indicated a future interest in self-tracking. (2) Desire for future tracking was linked to gaining a better understanding of whether any new symptoms that occur might be caused by MS.

5.2.2 Evaluation Interviews. We then recruited three survey respondents for interview and formative evaluation of our Fall Risk Monitor design scenario. To reflect our design scenario, we recruited participants with a relatively recent MS diagnosis and high levels of functional mobility. We met with each participant individually over Zoom. The design scenario was shown to participants as a presentation (see section 5.2.3). Sessions lasted around an hour and were recorded. We started each session by asking the participant to introduce themselves and say a little about how they experience living with MS and its symptoms. We then probed on responses they had made to our survey questions to gain deeper insight into their perspectives on self-tracking. Following this we played a video recording presenting the design scenario storyboard step-by-step with a voice-over explanation. We then asked the participant for their initial response and thoughts. Finally, we went through the storyboard presentation slide by slide probing more deeply on the participant's initial responses. Formative evaluation sessions were facilitated by the first author.

- *Participant PLWMS1* was aged twenty-seven, identified as female, and had been diagnosed with MS twelve months prior to the interview. She did not currently track symptoms, which she said present occasionally and are typically mild, but indicated an interest in trying out self-tracking in future.
- *Participant PLWMS2* was aged thirty-two, identified as female, and had been diagnosed with MS five years prior to the interview. She did not currently track symptoms, which she

said can be severe on occasion, and indicated that in most cases tracking is an unhelpful reminder of MS.
- *Participant PLWMS3* was aged thirty-four, identified as non-binary / third gender, and had been diagnosed with MS fifteen months prior to the interview. They did not currently track symptoms, which they said typically present as mild but can be frequent. They indicated an interest in trying self-tracking in future.

5.2.3 Fall Risk Monitor Design Scenario. We updated the Fall Risk Monitor design scenario previously presented to PTs in the speed dating sessions, and translated the storyboard format into presentation slides. Our updates included additional background information to the project, a brief introduction to wearables and AI, and additional detail explaining the scenario as it unfolds. The Fall Risk Monitor scenario was selected because it is more strongly focused on the perspective of someone who is living with a diagnosis of MS than the APTA scenario, and because PTs had highlighted potential drawbacks in the way video was used in the TAMPA scenario (see 4.3.6 for details). In addition, it reflected concerns that were visible in our analysis of the Reddit posts, such as understanding fatigue and how this might impact MS symptoms on 'bad days', and concern over future fall risk as a motivation for using a wearable. We created a video presentation to show participants online.

5.2.4 Analysis. We re-watched the video recordings multiple times, and amended the automated transcripts as necessary. We then conducted an inductive thematic analysis [13] of the transcripts from each participant together, to identify cross-cutting patterns and themes. The first and second authors each conducted an initial analysis individually, before coming together for agreement reaching, with the last author providing confidence in the derived themes. This resulted in eight additional findings and three framing concepts.

5.2.5 Findings. (1) Self-tracking should support better understanding of MS, and thereby agency over its impacts. For example, differentiating flares from pseudo-flares, or identifying which new symptoms are related to MS and which not. People have different symptoms and not all are caused by MS. (2) Wearables that automate data collection are appealing, but mood and fatigue prompts could become intrusive if they're not sensitive to individual preferences or cognitive issues e.g., 'brain fog', that can impact concentration. (3) Support understanding of everything that is happening around a particular behavior by tracking trends and their influences. For example, how does time of day or ambient temperature affect tiredness or bad mood? (4) It can be difficult to get a timely specialist appointment outside of emergencies, so to increase control and agency AI recommendations should help people make small changes independently without having to wait for an appointment. (5) Support for agency and autonomous change could also come from AI curated resources created by medical providers or verified and reputable experiential blogs. (6) Life continues outside of MS, and no noticeable change is often the most desirable outcome of a clinical appointment. (7) Focusing on personally meaningful goals and activities can help motivation, while collecting data all the time may highlight unexpected patterns. (8) Conducting research to benefit others can be a stronger motivation than self-tracking for personal benefit.

Illustrating our first finding, PLWMS3 highlighted challenges in their previous attempts at self-tracking saying, *"In the past, when I've tried to track symptoms and stuff, it seemed like the symptoms that I was experiencing or wanting to track, it was hard for me to describe them in a way that was useful to track. You know, things like my leg feels weird, you know. It's not really like quantitative. I can't put a 1 to 5 rating on it or anything. It's just things like that, and things like levels of fatigue. It's hard to know if I'm fatigued because of the MS, or if I'm fatigued because I had a hard day at work yesterday. Or if I had a hard day at work yesterday because I was experiencing MS fatigue"*. Illustrating the third finding, PLWMS2 said. *"So going back to like the temperature, or if it was windy that day, or the elevation, and what time of day they actually took a walk. Because I do much better in the morning than I do in the afternoon. So having that data be shared and having the AI be able to basically take that off the desk of a medical assistant or a neurologist. I think, would be helpful with like walking dogs. If I knew like, if the temperature goes above 85 you walk less, you walk weirder. And I personally get very irritated. I just get really pissy when I get too hot. So if I could have that data collected and analyzed by AI, and then it'd be like, Yeah, maybe go on your walks between these temperatures or these times. I think that would be helpful, not just helpful to improve mood. I'm not necessarily a fall risk, but that would help make my life a little better, without me having to analyze it, which is what I would do otherwise, and which would take forever"*.

5.2.6 Framing concepts: **(1)** ***'The personal mentor'*** *helps maintain focus on an individual's particular interests and goals; and supports them in doing what they're already doing, just better.* This concept reflects how the context of self-tracking is important and should reflect the goals and interests of PLWMS. **(2)** ***'The trouble shooter'*** *helps in understanding the present and preparing for the future; and supports independent change-making to help control the uncontrollable.* This reflects the challenges PLWMS can find in getting specialist help, and the opportunities that data and AI might offer to mitigate this. **(3)** ***'The research partner'*** *helps in finding and sharing knowledge and experiences, which contribute to better understanding and improved action.* This reflects the desire PLWMS may have to better understand their own experiences and symptoms, as well as to contribute towards better medical understanding of the condition itself.

6 SUMMARY AND SYNTHESIS

We bring together our insights in two storyboards (Figs 5 and 6), illustrating how our framing concepts (listed in Table 1) might be implemented. First we introduce Taylor (Fig 5), whose MS diagnosis is recent, along with their neurologist Charlie and primary care doctor Jesse. This storyboard illustrates how contextualized data from wearables might support Taylor's continued participation in her walking club. We use reference numbers to items in Table 1 to draw particular attention to how individual framing concepts have been illustrated, but the concepts are also used more holistically across these scenarios, and so this one-to-one linking is only a partial representation of their value. In Taylor's storyboard we draw attention to framing concepts ***2***, ***3***, ***5***, ***6***, and ***7***. Following this,

Table 1: Full listing of our framing concepts and the design activities they were derived from

Framing Concept	Design Activity
(1) ***'The ideal assistant'*** helps the clinician maintain a clear and present focus on their patient, and support them with timely information.	4.1 Understanding PT's current practice
(2) ***'The data interpreter'*** uses insights from outside the clinic to monitor change and help resolve tensions between short-term aims and longer-term consequences.	4.2 Exploring future practices
(3) ***'The facilitator'*** helps to open up patient-clinician conversations, scaffold decision-making, and prompt considerations that might otherwise be overlooked.	4.3 Co-creation with cards
(4) ***'The motivator'*** informs and encourages PLWMS, and supports adaptive responses to changing circumstances.	4.3 Co-creation with cards
(5) ***'The community-based assistant'*** helps bridge the gap between home and clinic, manage day-to-day interactions, and provide reliable and easy to digest insights and advice.	4.4 Speed dating with storyboards
(6) ***'The personal mentor'*** helps maintain focus on an individual's particular interests and goals; and supports them in doing what they're already doing, just better.	5 Learning from people living with MS
(7) ***'The trouble shooter'*** helps in understanding the present and preparing for the future; and supports independent change-making to help control the uncontrollable.	5 Learning from people living with MS
(8) ***'The research partner'*** helps in finding and sharing knowledge and experiences, which contribute to improved understanding and action.	5 Learning from people living with MS

we then introduce Sam (Fig 6), who has been living with MS for around 20 years, as well as their neurologic PT Alex, and reintroduce neurologist Charlie. Charlie is concerned that Sam's MS may be transitioning and wants to monitor for early signs of change. In Sam's storyboard we draw attention to framing concepts ***1***, ***2***, ***3***, ***4***, ***5***, ***6***, and ***8***.

7 DISCUSSION

We bring together our framing concepts in Table 1, and illustrate our synthesized findings in two storyboards (Figures: 5 and 6), illustrating how they might be implemented, and representing our current perspective on the knowledge generated during this research. Our findings extend prior literature by highlighting the importance of contextual information to translating data about gait metrics from wearable sensors into insight about mobility and community ambulation in the context of MS care. They also point towards ways in which this context might be framed and constrained to ease how contextual information is captured. In so doing, they suggest potential ways to ease previously highlighted tensions between clinical remote patient monitoring and agentic self-tracking.

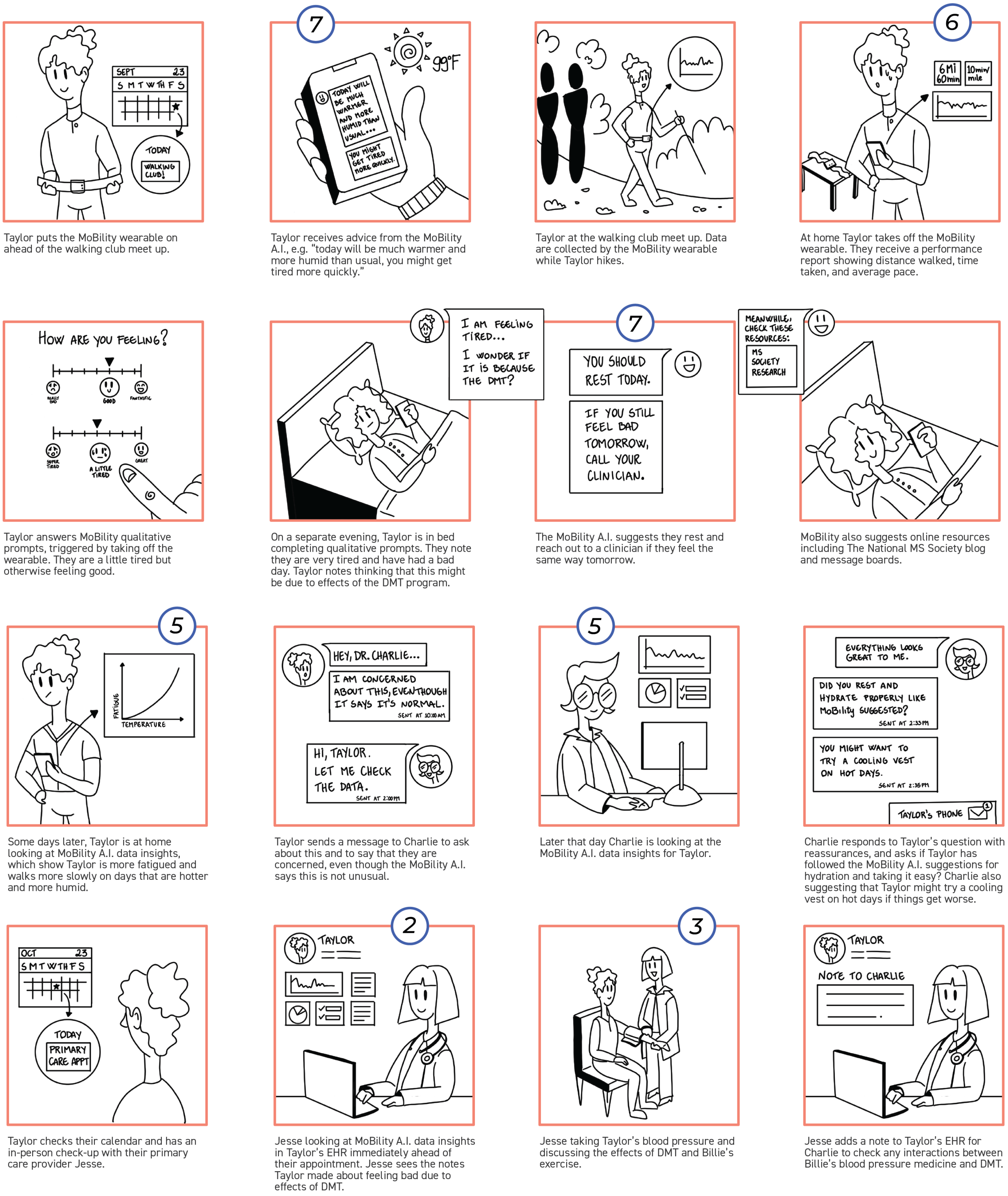
7.1 Translating Data About Gait Into Insight About Mobility

The PTs that participated in our co-design activities described themselves as 'talkers', highlighting the crucial role of patient-clinician conversations to their work. These conversations probe on things beyond what the PT is seeing, and try to draw out a patient's experiences away from the clinic. This provides contextual background that helps PTs to better understand and respond to standardized assessments, and provides insight into what is personally important to the patient. PTs' hopes for self-tracking data were associated with the potential to guide and prompt these conversations, and in particular to offer a window on the patient's community ambulation that might help in understanding the effectiveness of treatment interventions or patient experiences of pain and fatigue, or their trajectory towards achieving functional goals. Rehabilitation medicine is often focused on achieving or maintaining functional capabilities [76], and so PTs suggested self-tracking data could be particularly helpful for patients who struggle to perceive the benefits of an intervention. Comparing data from different time points, and to normative data or rates of expected change, could help where effects are slow and gradual, or where the goal is to slow the rate of decline as much as possible. As the quote from PT3 that we use in our title indicates, the gait metrics alone are of limited value, and the PTs want to know as much as possible about the context they were generated in. Prior work investigating cardiovascular healthcare [3] suggests that significant time, effort and care will be needed to support patients in interpreting and making sense of data, so that they can participate fully in collaborative processes. While better understanding of the sensemaking practices that will support patients [48, 49] and clinicians [70] will help, our findings in section 4.3.4 also highlight another key insight. This was the reminder that all data are not equal, and that the social context in which the incidents they represent took place may add significant weight to their perceived importance to the PLWMS. Insights from longitudinal self-tracking data will likely change these clinician-patient conversational practices, as was found in practices associated with patient generated data in oncology, which became more closely focused on validation [35]. A focus on validating aggregated data insights might downplay a single incident as atypical, and so designers should take care to avoid closing off opportunities for important conversation topics when thinking about sensemaking tools.

The PLWMS who participated in our formative evaluations were keen to stress that life continues outside of MS and that not all change is due to the impact of MS, but also that it can be hard to differentiate between a flare in MS symptoms or a pseudo-flare. In this respect, understanding the context of change can be extremely valuable e.g., in helping to unpick which new symptoms are related to MS and which not. For example, participant PLWMS2 felt

Taylor's storyboard

***Background and intro:** Taylor has recently been diagnosed with MS, their symptoms include recurring fatigue, and occasional numbness to their left side. Taylor is a keen hiker and member of a walking club. Taylor is concerned that people living with MS can often experience debilitating walking impairments, and so their neurologist Charlie has prescribed MoBility for self-tracking, alongside a disease modifying therapy (DMT) program. Insights from MoBility are also made available to Jesse, Taylor's primary care provider.*



Taylor puts the MoBility wearable on ahead of the walking club meet up.

Taylor receives advice from the MoBility A.I., e.g. "today will be much warmer and more humid than usual, you might get tired more quickly."

Taylor at the walking club meet up. Data are collected by the MoBility wearable while Taylor hikes.

At home Taylor takes off the MoBility wearable. They receive a performance report showing distance walked, time taken, and average pace.

Taylor answers MoBility qualitative prompts, triggered by taking off the wearable. They are a little tired but otherwise feeling good.

On a separate evening, Taylor is in bed completing qualitative prompts. They note they are very tired and have had a bad day. Taylor notes thinking that this might be due to effects of the DMT program.

The MoBility A.I. suggests they rest and reach out to a clinician if they feel the same way tomorrow.

MoBility also suggests online resources including The National MS Society blog and message boards.

Some days later, Taylor is at home looking at MoBility A.I. data insights, which show Taylor is more fatigued and walks more slowly on days that are hotter and more humid.

Taylor sends a message to Charlie to ask about this and to say that they are concerned, even though the MoBility A.I. says this is not unusual.

Later that day Charlie is looking at the MoBility A.I. data insights for Taylor.

Charlie responds to Taylor's question with reassurances, and asks if Taylor has followed the MoBility A.I. suggestions for hydration and taking it easy? Charlie also suggesting that Taylor might try a cooling vest on hot days if things get worse.

Taylor checks their calendar and has an in-person check-up with their primary care provider Jesse.

Jesse looking at MoBility A.I. data insights in Taylor's EHR immediately ahead of their appointment. Jesse sees the notes Taylor made about feeling bad due to effects of DMT.

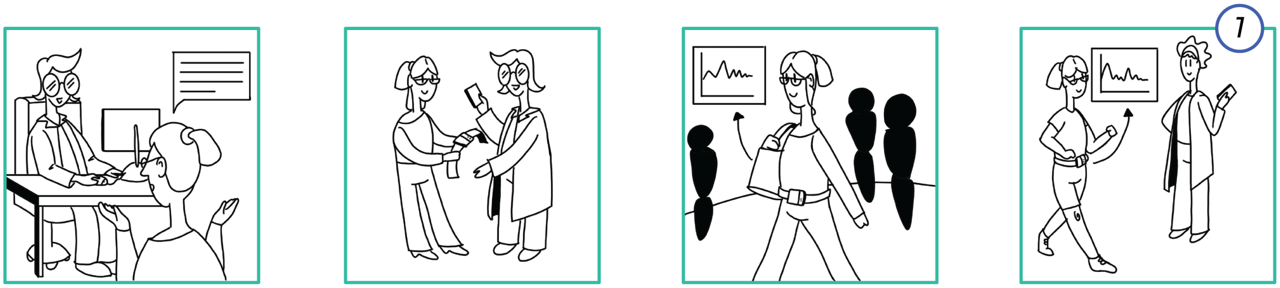
Jesse taking Taylor's blood pressure and discussing the effects of DMT and Billie's exercise.

Jesse adds a note to Taylor's EHR for Charlie to check any interactions between Billie's blood pressure medicine and DMT.

Figure 5: Taylor's storyboard, a design scenario illustrating how a person who has recently received a diagnosis of MS and their clinicians might use the MoBility self-tracking wearable and A.I. system. References to particular framing concepts are circled.

Sam’s storyboard

Background and intro: *Sam has been living with MS for around 20 years. Over that time, they have had occasional severe and debilitating flares. Recently these have become more frequent. Charlie, Sam’s neurologist, is concerned this may be an early sign of transitioning into secondary progressive MS. Charlie prescribes using MoBility while doing everyday activities, alongside a course of specialist physical therapy. Insights from MoBility A.I. will be shared with Alex the neurologic physical therapist.*

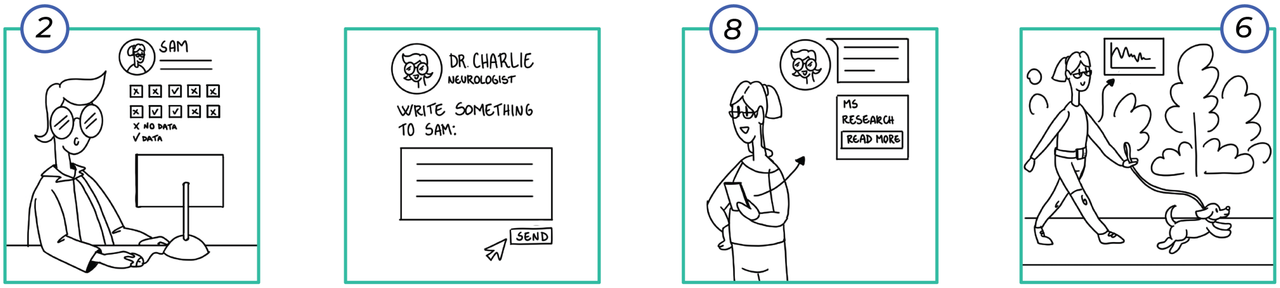


Sam is at their regular 6-month appointment with neurologist Charlie. They discuss how Sam’s flares are now more frequent and walking issues more common.

Charlie prescribes MoBility, shows Sam how to wear the device and complete the prompts, and explains that MoBility data updates and insights will be sent weekly for Charlie to review.

Sam practices using the MoBility device on a short trip to the local store.

Sam is at their physical therapy assessment with Alex. MoBility is used to capture gait data during 25-foot timed-walk test.

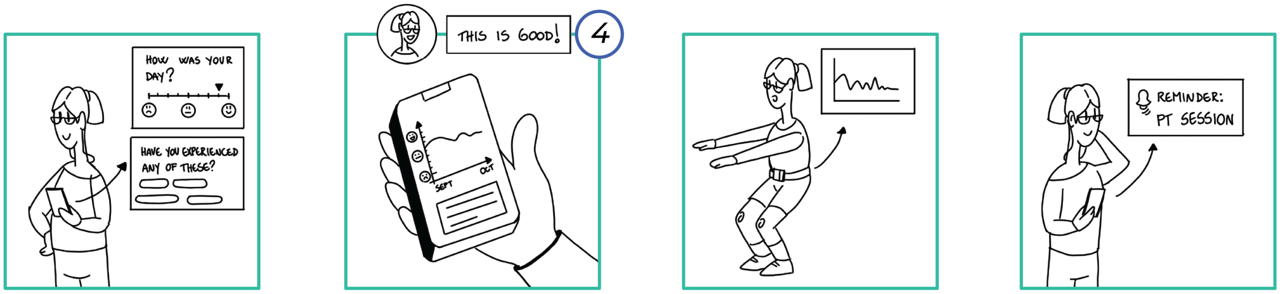


Charlie reviews Sam’s weekly insights briefing that MoBility’s A.I. has added to their electronic health record (EHR) and sees that Sam has been sporadic in using the device.

Charlie adds an encouragement and reminder prompt into Sam’s MoBility A.I.

Sam sees the reminder set by Charlie, which includes a link thet reminds Sam they are helping MS research.

Sam focuses on wearing the MoBility device when they are walking their dog.

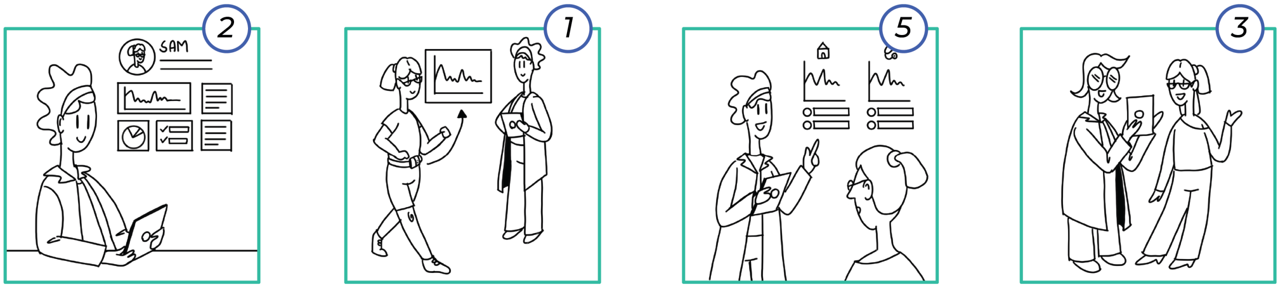


Sam completes daily qualitative prompts each evening. Today has been just a regular day with no drama.

Sam sees that most days are just regular, no drama days. This confirms that Sam is generally staying well.

Sam wears the MoBility device to collect data as they complete the at-home exercise program Alex has set for them.

Later on, Sam gets a reminder that they are due for a follow-up physical therapy session with Alex the following day.



Alex is viewing Sam’s MoBility data and A.I. generated insights ahead of their follow-up appointment.

Sam’s follow-up physical therapy appointment. They wear the MoBility device to complete the exercise program with Alex.

Alex shows Sam the MoBility insights from at home and in the clinic, and they discuss an updated exercise program.

It is Sam’s next 6-month appointment with Charlie. They are both looking at the MoBility data and A.I. insights, and discussing how the program is going.

Figure 6: Sam’s storyboard, a design scenario illustrating how a person who has been living with MS for a number of years might use the MoBility technology with their clinicians. References to particular framing concepts are circled.

that with a better understanding of "everything that's happening around" a particular activity, they could have greater agency in self-selecting mitigation actions. This resonates with prior research that highlights the importance of personal relevance in selecting self-tracking metrics [7], and how these connect to clinical measures and objectives [83]. However, as we discuss below, framing and constraining the "everything that's happening around" is a key challenge for future research. As an example, participant PLWMS2 discussed how heat can cause fatigue that impairs their walking and makes them irritable, and that being irritable can make it harder to make the best choices at the time. Having data that would allow them to connect different factors and see patterns behind the behaviors could help in identifying and understanding the trigger, and prompt them to take preemptive measures to help them stay cool. We see opportunities for self-tracking practices like those illustrated in Taylor's storyboard (Fig. 5) to indicate possible patterns connecting fatigue, walking impairment, and bad mood. However, we might go further and add data sources to broaden the picture with additional context, such as ambient temperature and weather, and use the resulting insights to offer advice such as wearing a cooling vest. In this way, we see opportunities to support the everyday self-care called for in prior research [52, 53].

7.2 Framing and Constraining Context

Our research highlights the importance of context to understanding and making sense of gait metrics, in order that they support actionable insight about mobility for PLWMS and their clinicians. However, the concept of context can itself be elusive [22], and contextual uncertainty can confound these processes [5]. Because of this, we suggest that placing constraints around what is considered, and reducing the potential for contextual ambiguity will help PLWMS and their clinicians gain understanding and insight. Our findings highlight the potential for achieving this by focusing the situations in which gait metrics data are generated on the functional goals agreed between PTs and PLWMS; and in particular through selecting and focusing on a specific activity that is an important motivator for the PLWMS maintaining mobility. This aligns well with rehabilitative care's focus on capacity[76] as well as with the priority concerns discussed by PLWMS, and extends prior work on self-trcking motivations. PTs and PLWMS both drew attention to the potential for this approach to help maintain motivation, limit concerns about privacy, and reduce the likelihood that prompts will seem intrusive. It also facilitates more straightforward, like-for-like comparison, and can support PTs and PLWMS in planning individual protocols for selecting, collecting, and making sense of data. For example, one situation mentioned was walking the dog. Some people vary where this takes take place, and so capturing GPS metadata for location and route information could support insight into weather conditions [37] and terrain [30], and thereby support greater understanding of the contexts in which mobility is adversely affected. For others, dog walking may be repetitive and follow the same routine in the same location, making a GPS record less important to understanding context, as this aspect is relatively constant. Similarly, the social context of walking might also vary in important ways. For example, when collecting children from school it may be important to know whether the PLWMS is alone or in a group with multiple children and adults, as this is likely reflected in patterns of gait metric data but not explicitly identified. We believe that capturing information such as this could become part of experience sampling protocols [18, 78], and that these protocols could also probe on factors such as mood, fatigue, and pain, which also impact mobility. Capturing this information as it happens offers another key benefit, as it means that PLWMS do not have to try to remember what was happening on different days, when their clinician is probing on a pattern in gait metric data.

7.3 Using AI to Reduce the Burden on Clinicians and People Living with MS

While data can help in understanding important contextual factors, care will need to be taken as the introduction of data work into healthcare places additional burdens on all concerned. Research has shown that virtual first, continuous and connected care [50], such as patients' self-tracking, creates additional work for clinicians [9, 42] that results in digital tasks being completed outside of working hours [41, 65], and that high volumes of data may discourage widespread clinical adoption of new digital practices [51]. Also, that for PLWMS, self-tracking can feel like burdensome work or provide an unnecessary reminder of MS [51, 82]. With this in mind we introduced AI into our design scenarios to explore how routine data analyses and straightforward patient-clinician interactions might be supported through additional automation. Initially, we had some concerns that AI may not seem relevant or may confuse participants. However, we found that the PTs and PLWMS all took the potential use of AI in their stride. For example, outpatient PTs noted how they prepare for appointments right before the patient arrives, and thought that AI highlights would be a useful synopsis and conversation prompt; while for PLWMS, arranging timely appointments with specialists can be a major challenge and AI was seen as a way to help mitigate this. With this in mind, PLWMS1 suggested that AI might monitor patterns in their data and propose simple mitigation options for low-risk concerns that could be tried without the need for consultation with a clinician. For example, suggesting that they try a cooling jacket if heat appears to correlate with fatigue that negatively impacts their mobility. If the suggested options do not improve or resolve the situation, their clinician could be contacted; and where suggestions did help, the clinician would have visibility and be more informed and better able to help should the issue return. Moving forward, inquiry should address the level of risk appropriate in deciding on AI autonomy, particularly with regard to how accurate and reliable predictions need to be to have value. For example, is a self-tracking recommendation more like route planning in Google Maps, in that it will suggest a best option with only one or two close alternatives, or more like ranked-choice in Amazon or Netflix? And are there similarities between how a Nest learns patterns of activity to automate changes in heating and cooling? Prior work that offers guidance in how to approach these concerns, includes: support for human-AI collaboration in mixed initiative settings [33]; the use of metaphor to identify assumptions with regard to what is known and what unknown in designing with AI [23]; and on how co-design ideation with AI can consider feasibility as well as desirability [85]. In the context of longitudinal care for a chronic debilitating condition, it is also necessary to remain open to ways motivations and configurations might vary over time, as situations and symptoms change.

7.4 Easing Tensions Between Clinical Remote Patient Monitoring, and Agentic Self-Tracking and Care

A key aspect highlighted by previous research, in particular Nunes et al. [54], is a tension between patient autonomy and clinician control, where devices are focused either on medical symptoms or the patient's lived experience. This reflects wider tensions in healthcare self-monitoring that can be seen in the contrast between strong clinical concerns about adherence [42, 58] that are critiqued in HCI [52]. We believe that this emerges from and reflects different disciplinary motivations and priorities. For example, HCI research is typically conducted with self-selecting and motivated participants e.g., [4, 8, 53], while clinical self-monitoring studies are rooted in at-home hypertension and diabetes monitoring and has typically included those who can struggle to participate fully for any numbe of technical, financial and social reasons [17]. Our findings suggests these tensions can be eased in at least two ways. First, focusing on functional goals and priority activities helps because it builds on shared objectives and brings divergent motivations closer together. In this way it can also help mitigate situations where PLWMS find the activity of self-tracking to be burdensome work or an unnecessary reminder of MS [51, 82]. Our evaluations with PLWMS also highlighted important nuances in how we might think about this. For example, MS can be a cause of changing mood and the cognitive impairment PLWMS3 described as 'brain fog', which can increase the cognitive demands of self-tracking. However, while self-tracking might sometimes seem annoying or intrusive, timely help in understanding the cause of new symptoms, distinguishing flares from pseudo-flares, and planning mitigation tactics, can also provide important help and be a reason to complete experience sampling prompts or remember a wearable. We illustrate this in our design scenario in Fig 5, which shows MoBility notifying Taylor of the potential for raised temperatures as they are preparing to go to their walking club meet up, and later facilitating self-care in response. Second, contributing to research can also be a strong motivation. PLWMS2 suggested that benefiting others would be more likely to motivate continued participation than self-tracking for its own sake. A similar finding is reported in [82], and we reflect it in our design scenario Fig 6. However, this is a design opportunity that has yet to be explored in depth. PLWMS2 compared it to posting on subreddits, where people commonly join to ask questions, but over time transition into a source of knowledge and reassurance, and suggested that providing reminders of the ongoing and long-term results from research, at each instance of self-tracking, would be a great opportunity to feel like "I got to see something good is happening".

7.5 Limitations and Future Work

While our research provides initial pointers to how these tensions might be eased, we also acknowledge that our design process was skewed towards the perspective of clinicians. We undertook four rounds of design activities with PTs and just a single round of formative evaluations with PLWMS. These choices have consequences, and our framing concepts were primarily informed by the needs and desires of clinicians. The first five of these emerge from our interactions with PTs, and share a common theme of helping clinicians gain insight into the experiences patients have outside of clinical appointments. In our design scenarios this manifests, for example, in our including experience sampling prompts that help to contextualize data about gait metrics for the benefit of PTs' understanding. While the final three framing concepts emerge from interactions with PLWMS, two of these focus broadly on supporting adherence (*The personal mentor* and *The research partner*), while only one (*The trouble shooter*) aims at supporting similar insight and understanding for PLWMS that we offer PTs. This manifests, for example, in our design scenarios requiring PLWMS to make appointments with clinicians when their own understanding of gait metrics data requires a detailed explanation. While this reflects current power and knowledge relations in medicine, it is not a necessity from a design perspective. In our ongoing work, which will be reported in future research papers, we are starting from the perspective of PLWMS and conducting design inquiry to explore competing and complimentary needs and desires. This will result in a different set of framing concepts. Making things explicit in this way will support reflection on design judgements and trade-offs when bridging the perspectives of clinicians and PLWMS.

Beyond this, there are three main limitations to be acknowledged. First, this exploratory design-focused study engaged with a limited number of practitioners, focusing on PTs in a single medical center system. This is mitigated by our longitudinal approach, which allows for deeper engagement with PTs than interview alone, and by revisiting data from our prior studies. However, hospitals and care sites differ in complex ways, as does the work of healthcare practitioners, with significant implications for transferability. Second, this research engages with only a small number of PLWMS, and recruitment was restricted to Reddit, further limiting scope and reach. Our future research will probe a broader range of experiences e.g., by recruiting through Facebook groups and the National MS Society message boards, and also consider carers and others who may be impacted. Third, our design iterations have so far been been conceptual. While this allows for valuable ambiguity in the early stages of design, different stages of prototyping introduce different challenges and opportunities. Moving forward, our work will progress through iterations of exploratory prototyping at different levels of fidelity and with different technical capabilities implemented.

8 CONCLUSION

In this paper we present design inquiry into how self-tracking using wearable sensors might be incorporated into the healthcare and self-care of people living with multiple sclerosis (MS). We detail a multi-stage process that includes longitudinal inquiry with physical therapists (PTs), analysis of public Reddit forums, and formative evaluations with people living with MS (PLWMS). We present design scenarios and storyboards that illustrate how changes in gait, mobility, and balance, might be monitored to facilitate timely diagnosis of walking impairments, which are among the most common and debilitating symptoms of MS. We offer framing concepts to guide interaction design in this and similar contexts, and contribute to HCI understanding of self-tracking in rehabilitation healthcare.

Finally, we contribute to HCI understanding of virtual-first practices in rehabilitation medicine that can be translated for use-cases beyond MS care.

ACKNOWLEDGMENTS

We'd like to thank the physical therapists and people living with multiple sclerosis whose participation made this research possible. This research was supported by NSF awards #1928614 and #2129076.

REFERENCES

[1] Kevin N Alschuler and Meghan L Beier. 2015. Intolerance of uncertainty: shaping an agenda for research on coping with multiple sclerosis. *International journal of MS care* 17, 4 (2015), 153–158.

[2] Ashley Amaya, Ruben Bach, Florian Keusch, and Frauke Kreuter. 2021. New data sources in social science research: Things to know before working with Reddit data. *Social science computer review* 39, 5 (2021), 943–960.

[3] Tariq Andersen, Pernille Bjørn, Finn Kensing, and Jonas Moll. 2011. Designing for collaborative interpretation in telemonitoring: Re-introducing patients as diagnostic agents. *International journal of medical informatics* 80, 8 (2011), e112–e126.

[4] Tariq Osman Andersen, Jonas Fritsch, and Stina Matthiesen. 2022. Patient Data Work with Consumer Self-tracking: Exploring Affective and Temporal Dimensions in Chronic Self-care. In *International Conference on Pervasive Computing Technologies for Healthcare*. Springer, 666–680.

[5] Tariq Osman Andersen, Henriette Langstrup, and Stine Lomborg. 2020. Experiences with wearable activity data during self-care by chronic heart patients: qualitative study. *Journal of medical Internet research* 22, 7 (2020), e15873.

[6] Amid Ayobi. 2020. *Self-Tracking by People Living with Multiple Sclerosis: Supporting Experiences of Agency in a Chronic Neurological Condition.* Ph. D. Dissertation. UCL (University College London).

[7] Amid Ayobi, Paul Marshall, and Anna L Cox. 2020. Trackly: A Customisable and Pictorial Self-Tracking App to Support Agency in Multiple Sclerosis Self-Care. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.

[8] Amid Ayobi, Paul Marshall, Anna L Cox, and Yunan Chen. 2017. Quantifying the body and caring for the mind: self-tracking in multiple sclerosis. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 6889–6901.

[9] Jakob E Bardram, Claus Bossen, and Anders Thomsen. 2005. Designing for transformations in collaboration: a study of the deployment of homecare technology. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. 294–303.

[10] Jeffrey Bardzell, Shaowen Bardzell, Peter Dalsgaard, Shad Gross, and Kim Halskov. 2016. Documenting the research through design process. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. 96–107.

[11] Lauren C Benson, Christian A Clermont, Eva Bošnjak, and Reed Ferber. 2018. The use of wearable devices for walking and running gait analysis outside of the lab: A systematic review. *Gait & posture* 63 (2018), 124–138.

[12] Michael J Bradshaw, Samantha Farrow, Robert W Motl, and Tanuja Chitnis. 2017. Wearable biosensors to monitor disability in multiple sclerosis. *Neurology: Clinical Practice* 7, 4 (2017), 354–362.

[13] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[14] Wallace J Brownlee, Todd A Hardy, Franz Fazekas, and David H Miller. 2017. Diagnosis of multiple sclerosis: progress and challenges. *The Lancet* 389, 10076 (2017), 1336–1346.

[15] Michelle H Cameron and Joanne M Wagner. 2011. Gait abnormalities in multiple sclerosis: pathogenesis, evaluation, and advances in treatment. *Current neurology and neuroscience reports* 11 (2011), 507–515.

[16] Tommaso Ciceri, Giorgia Malerba, Alice Gatti, Eleonora Diella, Denis Peruzzo, Emilia Biffi, and Luca Casartelli. 2023. Context expectation influences the gait pattern biomechanics. *Scientific Reports* 13, 1 (2023), 5644.

[17] SF Clarke and JR Foster. 2012. A history of blood glucose meters and their role in self-monitoring of diabetes mellitus. *British journal of biomedical science* 69, 2 (2012), 83–93.

[18] Mihaly Csikszentmihalyi and Reed Larson. 1987. Validity and reliability of the experience-sampling method. *The Journal of nervous and mental disease* 175, 9 (1987), 526–536.

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

[20] Laura Dennison, Ellen McCloy Smith, Katherine Bradbury, and Ian Galea. 2016. How do people with multiple sclerosis experience prognostic uncertainty and prognosis communication? A qualitative study. *PloS one* 11, 7 (2016), e0158982.

[21] Kees Dorst. 2011. The core of 'design thinking'and its application. *Design studies* 32, 6 (2011), 521–532.

[22] Paul Dourish. 2004. What we talk about when we talk about context. *Personal and ubiquitous computing* 8 (2004), 19–30.

[23] Graham Dove and Anne-Laure Fayard. 2020. Monsters, metaphors, and machine learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–17.

[24] Graham Dove, Nicolai Brodersen Hansen, and Kim Halskov. 2016. An argument for design space reflection. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. 1–10.

[25] Graham Dove, Caroline Emilie Lundqvist, and Kim Halskov. 2018. The life cycle of a generative design metaphor. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*. ACM, ACM, New York, NY, 414–425.

[26] Valery L Feigin, Amanuel Alemu Abajobir, Kalkidan Hassen Abate, Foad Abd-Allah, Abdishakur M Abdulle, Semaw Ferede Abera, Gebre Yitayih Abyu, Muktar Beshir Ahmed, Amani Nidhal Aichour, Ibtihel Aichour, et al. 2017. Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet Neurology* 16, 11 (2017), 877–897.

[27] Felix Flachenecker, Heiko Gaßner, Julius Hannik, De-Hyung Lee, Peter Flachenecker, Jürgen Winkler, Bjoern Eskofier, Ralf A Linker, and Jochen Klucken. 2020. Objective sensor-based gait measures reflect motor impairment in multiple sclerosis patients: reliability and clinical validation of a wearable sensor device. *Multiple sclerosis and related disorders* 39 (2020), 101903.

[28] Douglas D Gunzler, Alessandro S De Nadai, Deborah M Miller, Daniel Ontaneda, and Farren BS Briggs. 2023. Long-term trajectories of ambulatory impairment in multiple sclerosis. *Multiple Sclerosis Journal* (2023), 13524585231187521.

[29] Kim Halskov and Peter Dalsgård. 2006. Inspiration card workshops. In *Proceedings of the 6th conference on Designing Interactive systems*. 2–11.

[30] Jordy Hendrikx, Jerry Johnson, and Christopher Shelly. 2016. Using GPS tracking to explore terrain preferences of heli-ski guides. *Journal of outdoor recreation and tourism* 13 (2016), 34–43.

[31] Laura Holsbeeke, Marjolijn Ketelaar, Marina M Schoemaker, and Jan Willem Gorter. 2009. Capacity, capability, and performance: different constructs or three of a kind? *Archives of physical medicine and rehabilitation* 90, 5 (2009), 849–855.

[32] Karen Holtzblatt and Hugh Beyer. 2017. *Contextual Design: Design for Life* (second edition. ed.). Morgan Kaufmann.

[33] Eric Horvitz. 1999. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 159–166.

[34] Yonatan Hutabarat, Dai Owaki, and Mitsuhiro Hayashibe. 2021. Recent advances in quantitative gait analysis using wearable sensors: a review. *IEEE Sensors Journal* 21, 23 (2021), 26470–26487.

[35] Anna Sigridur Islind, Tomas Lindroth, Johan Lundin, and Gunnar Steineck. 2019. Shift in translations: Data work with patient-generated health data in clinical practice. *Health informatics journal* 25, 3 (2019), 577–586.

[36] Alon Kalron, Dalia Nitzani, David Magalashvili, Mark Dolev, Shay Menascu, Yael Stern, Uri Rosenblum, Diana Pasitselsky, Lior Frid, Gabi Zeilig, et al. 2015. A personalized, intense physical rehabilitation program improves walking in people with multiple sclerosis presenting with different levels of disability: a retrospective cohort. *BMC neurology* 15, 1 (2015), 1–9.

[37] Athar Kamal, Syed Mustafa Husain Abidi, Ahmed Mahfouz, Sambhaji Kadam, Aziz Rahman, Ibrahim Galal Hassan, and Liangzhou Leon Wang. 2021. Impact of urban morphology on urban microclimate and building energy loads. *Energy and Buildings* 253 (2021), 111499.

[38] Ilana Katz Sand, Stephen Krieger, Colleen Farrell, and Aaron E Miller. 2014. Diagnostic uncertainty during the transition to secondary progressive multiple sclerosis. *Multiple sclerosis journal* 20, 12 (2014), 1654–1657.

[39] Klaus Krippendorff. 1989. On the essential contexts of artifacts or on the proposition that" design is making sense (of things)". *Design issues* 5, 2 (1989), 9–39.

[40] Nicholas G LaRocca. 2011. Impact of walking impairment in multiple sclerosis. *The Patient: Patient-Centered Outcomes Research* 4, 3 (2011), 189–201.

[41] Katharine Lawrence, Oded Nov, Devin Mann, Soumik Mandal, Eduardo Iturrate, and Batia Wiesenfeld. 2022. The Impact of Telemedicine on Physicians' After-hours Electronic Health Record "Work Outside Work" During the COVID-19 Pandemic: Retrospective Cohort Study. *JMIR Medical Informatics* 10, 7 (2022), e34826.

[42] Katharine Lawrence, Nina Singh, Zoe Jonassen, Lisa L Groom, Veronica Alfaro Arias, Soumik Mandal, Antoinette Schoenthaler, Devin Mann, Oded Nov, and Graham Dove. 2023. Operational Implementation of Remote Patient Monitoring Within a Large Ambulatory Health System: Multimethod Qualitative Case Study. *JMIR Human Factors* 10, 1 (2023), e45166.

[43] Bryan Lawson and Shee Ming Loke. 1997. Computers, words and pictures. *Design studies* 18, 2 (1997), 171–183.

[44] Izabella Lejbkowicz, Tamar Paperna, Nili Stein, Sara Dishon, Ariel Miller, et al. 2010. Internet usage by patients with multiple sclerosis: implications to participatory medicine and personalized healthcare. *Multiple sclerosis international* 2010 (2010).

[45] SE Lord, PW Halligan, and DT Wade. 1998. Visual gait analysis: the development of a clinical assessment and scale. *Clinical rehabilitation* 12, 2 (1998), 107–119.

[46] Susan E Lord, Kathryn McPherson, Harry K McNaughton, Lynn Rochester, and Mark Weatherall. 2004. Community ambulation after stroke: how important and obtainable is it and what measures appear predictive? *Archives of physical medicine and rehabilitation* 85, 2 (2004), 234–239.

[47] Andrés Lucero, Peter Dalsgaard, Kim Halskov, and Jacob Buur. 2016. Designing with cards. *Collaboration in creative design: Methods and tools* (2016), 75–95.

[48] Lena Mamykina, Elizabeth Mynatt, Patricia Davidson, and Daniel Greenblatt. 2008. MAHI: investigation of social scaffolding for reflective thinking in diabetes management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 477–486.

[49] Lena Mamykina, Arlene M Smaldone, and Suzanne R Bakken. 2015. Adopting the sensemaking perspective for chronic disease self-management. *Journal of biomedical informatics* 56 (2015), 406–417.

[50] Devin M Mann and Katharine Lawrence. 2022. Reimagining connected care in the era of digital medicine. *JMIR mHealth and uHealth* 10, 4 (2022), e34483.

[51] Martin Marziniak, Giampaolo Brichetto, Peter Feys, Uta Meyding-Lamadé, Karen Vernon, Sven G Meuth, et al. 2018. The use of digital and remote communication technologies as a tool for multiple sclerosis management: narrative review. *JMIR rehabilitation and assistive technologies* 5, 1 (2018), e7805.

[52] Francisco Nunes. 2019. From medicalized to mundane self-care technologies. *interactions* 26, 3 (2019), 67–69.

[53] Francisco Nunes, Tariq Andersen, and Geraldine Fitzpatrick. 2019. The agency of patients and carers in medical care and self-care technologies for interacting with doctors. *Health informatics journal* 25, 2 (2019), 330–349.

[54] Francisco Nunes, Nervo Verdezoto, Geraldine Fitzpatrick, Morten Kyng, Erik Grönvall, and Cristiano Storni. 2015. Self-care technologies in HCI: Trends, tensions, and opportunities. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, 6 (2015), 1–45.

[55] Jiwon Oh, Katayoun Alikhani, Tania Bruno, Virginia Devonshire, Paul S Giacomini, Fabrizio Giuliani, Hamid R Nakhaipour, Robyn Schecter, and Catherine Larochelle. 2019. Diagnosis and management of secondary-progressive multiple sclerosis: time for change. *Neurodegenerative disease management* 9, 6 (2019), 301–317.

[56] Massimiliano Pau, Silvia Caggiari, Alessandro Mura, Federica Corona, Bruno Leban, Giancarlo Coghe, Lorena Lorefice, Maria Giovanna Marrosu, and Eleonora Cocco. 2016. Clinical assessment of gait in individuals with multiple sclerosis using wearable inertial sensors: Comparison with patient-based measure. *Multiple sclerosis and related disorders* 10 (2016), 187–191.

[57] Massimiliano Pau, Serena Mandaresu, Giuseppina Pilloni, Micaela Porta, Giancarlo Coghe, Maria Giovanna Marrosu, and Eleonora Cocco. 2017. Smoothness of gait detects early alterations of walking in persons with multiple sclerosis without disability. *Gait & posture* 58 (2017), 307–309.

[58] Joshua M Pevnick, Garth Fuller, Ray Duncan, and Brennan MR Spiegel. 2016. A large-scale initiative inviting patients to share personal fitness tracker data with their providers: initial results. *PloS one* 11, 11 (2016), e0165908.

[59] Ashley Polhemus, Laura Delgado Ortiz, Gavin Brittain, Nikolaos Chynkiamis, Francesca Salis, Heiko Gaßner, Michaela Gross, Cameron Kirk, Rachele Rossanigo, Kristin Taraldsen, et al. 2021. Walking on common ground: a cross-disciplinary scoping review on the clinical utility of digital mobility outcomes. *NPJ digital medicine* 4, 1 (2021), 149.

[60] Micaela Porta, Giuseppina Pilloni, Roberta Pili, Carlo Casula, Mauro Murgia, Giovanni Cossu, and Massimiliano Pau. 2018. Association between objectively measured physical activity and gait patterns in people with Parkinson's disease: results from a 3-month monitoring. *Parkinson's Disease* 2018 (2018).

[61] William L Porter. 1988. Notes on the inner logic of designing: Two thought-experiments. *Design Studies* 9, 3 (1988), 169–180.

[62] Nicholas Proferes, Naiyan Jones, Sarah Gilbert, Casey Fiesler, and Michael Zimmer. 2021. Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media+ Society* 7, 2 (2021), 20563051211019004.

[63] Nils Roth, Georg P Wieland, Arne Küderle, Martin Ullrich, Till Gladow, Franz Marxreiter, Jochen Klucken, Bjoern M Eskofier, and Felix Kluge. 2021. Do we walk differently at home? a context-aware gait analysis system in continuous real-world environments. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 1932–1935.

[64] Dalia Rotstein and Xavier Montalban. 2019. Reaching an evidence-based prognosis for personalized treatment of multiple sclerosis. *Nature Reviews Neurology* 15, 5 (2019), 287–300.

[65] Harry S Saag, Kanan Shah, Simon A Jones, Paul A Testa, and Leora I Horwitz. 2019. Pajama time: working after work in the electronic health record. *Journal of general internal medicine* 34, 9 (2019), 1695–1696.

[66] Rosaria Sacco, Rita Bussman, Peter Oesch, Jürg Kesselring, and Serafin Beer. 2011. Assessment of gait parameters and fatigue in MS patients during inpatient rehabilitation: a pilot trial. *Journal of neurology* 258, 5 (2011), 889–894.

[67] Elizabeth B-N Sanders and Pieter Jan Stappers. 2008. Co-creation and the new landscapes of design. *Co-design* 4, 1 (2008), 5–18.

[68] Donald A Schön. 1988. Designing: Rules, types and words. *Design studies* 9, 3 (1988), 181–190.

[69] Donald A. Schön. 1993. Generative metaphor: A perspective on problem-setting in social policy. In *In Metaphor and Thought (2nd ed.)*, Andrew Ortony (Ed.). Cambridge University Press, Cambridge, UK, 137–163.

[70] Ayanna Seals, Giuseppina Pilloni, Jin Kim, Raul Sanchez, John-Ross Rizzo, Leigh Charvet, Oded Nov, and Graham Dove. 2022. 'Are They Doing Better In The Clinic Or At Home?': Understanding Clinicians' Needs When Visualizing Wearable Sensor Data Used In Remote Gait Assessments For People With Multiple Sclerosis. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.

[71] Itamar Shatz. 2017. Fast, free, and targeted: Reddit as a source for recruiting participants online. *Social Science Computer Review* 35, 4 (2017), 537–549.

[72] Herbert A. (Herbert Alexander) Simon. 1969. *The Sciences of the Artificial.* M.I.T. Press, Cambridge.

[73] Alexandra Simpson, Ellen M Mowry, and Scott D Newsome. 2021. Early aggressive treatment approaches for multiple sclerosis. *Current treatment options in neurology* 23 (2021), 1–21.

[74] Charalampos Sotirakis, Zi Su, Maksymilian A Brzezicki, Niall Conway, Lionel Tarassenko, James J FitzGerald, and Chrystalina A Antoniades. 2023. Identification of motor progression in Parkinson's disease using wearable sensors and machine learning. *npj Parkinson's Disease* 9, 1 (2023), 142.

[75] Marc Steen. 2013. Co-design as a process of joint inquiry and imagination. *Design issues* 29, 2 (2013), 16–28.

[76] Gerold Stucki, Thomas Ewert, and Alarcos Cieza. 2002. Value and application of the ICF in rehabilitation medicine. *Disability and rehabilitation* 24, 17 (2002), 932–938.

[77] Lena H Ting, Hillel J Chiel, Randy D Trumbower, Jessica L Allen, J Lucas McKay, Madeleine E Hackney, and Trisha M Kesar. 2015. Neuromechanical principles underlying movement modularity and their implications for rehabilitation. *Neuron* 86, 1 (2015), 38–54.

[78] Niels Van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The experience sampling method on mobile devices. *ACM Computing Surveys (CSUR)* 50, 6 (2017), 1–40.

[79] Aliénor Vienne, Rémi P Barrois, Stéphane Buffat, Damien Ricard, and Pierre-Paul Vidal. 2017. Inertial sensors to assess gait quality in patients with neurological disorders: a systematic review of technical and analytical challenges. *Frontiers in psychology* 8 (2017), 817.

[80] Mitchell T Wallin, William J Culpepper, Jonathan D Campbell, Lorene M Nelson, Annette Langer-Gould, Ruth Ann Marrie, Gary R Cutter, Wendy E Kaye, Laurie Wagner, Helen Tremlett, et al. 2019. The prevalence of MS in the United States: a population-based estimate using health claims data. *Neurology* 92, 10 (2019), e1029–e1040.

[81] Clare Walton, Rachel King, Lindsay Rechtman, Wendy Kaye, Emmanuelle Leray, Ruth Ann Marrie, Neil Robertson, Nicholas La Rocca, Bernard Uitdehaag, Ingrid van der Mei, et al. 2020. Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS. *Multiple Sclerosis Journal* 26, 14 (2020), 1816–1821.

[82] Karine Wendrich and Lotte Krabbenborg. 2023. Negotiating with digital self-monitoring: A qualitative study on how patients with multiple sclerosis use and experience digital self-monitoring within a scientific study. *Health* (2023), 13634593231175321.

[83] Karine Wendrich, Pim van Oirschot, Marijn Bart Martens, Marco Heerings, Peter Joseph Jongen, and Lotte Krabbenborg. 2019. Toward digital self-monitoring of multiple sclerosis: investigating first experiences, needs, and wishes of people with MS. *International journal of MS care* 21, 6 (2019), 282–291.

[84] Tim Woelfle, Lucie Bourguignon, Johannes Lorscheider, Ludwig Kappos, Yvonne Naegelin, and Catherine Ruth Jutzeler. 2023. Wearable sensor technologies to assess motor functions in people with multiple sclerosis: systematic scoping review and perspective. *Journal of Medical Internet Research* 25 (2023), e44428.

[85] Nur Yildirim, Changhoon Oh, Deniz Sayar, Kayla Brand, Supritha Challa, Violet Turri, Nina Crosby Walton, Anna Elise Wong, Jodi Forlizzi, James McCann, and John Zimmerman. 2023. Creating Design Resources to Scaffold the Ideation of AI Concepts. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 2326–2346. https://doi.org/10.1145/3563657.3596058

[86] John Zimmerman and Jodi Forlizzi. 2014. Research through design in HCI. *Ways of Knowing in HCI* (2014), 167–189.

[87] John Zimmerman and Jodi Forlizzi. 2017. Speed dating: providing a menu of possible futures. *She Ji: The Journal of Design, Economics, and Innovation* 3, 1 (2017), 30–50.

[88] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 493–502.

[89] John Zimmerman, Aaron Steinfeld, Anthony Tomasic, and Oscar J. Romero. 2022. Recentering Reframing as an RtD Contribution: The Case of Pivoting from Accessible Web Tables to a Conversational Internet. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–14.

A APPENDIX 1: DIGITAL HEALTHCARE DESIGN CARDS

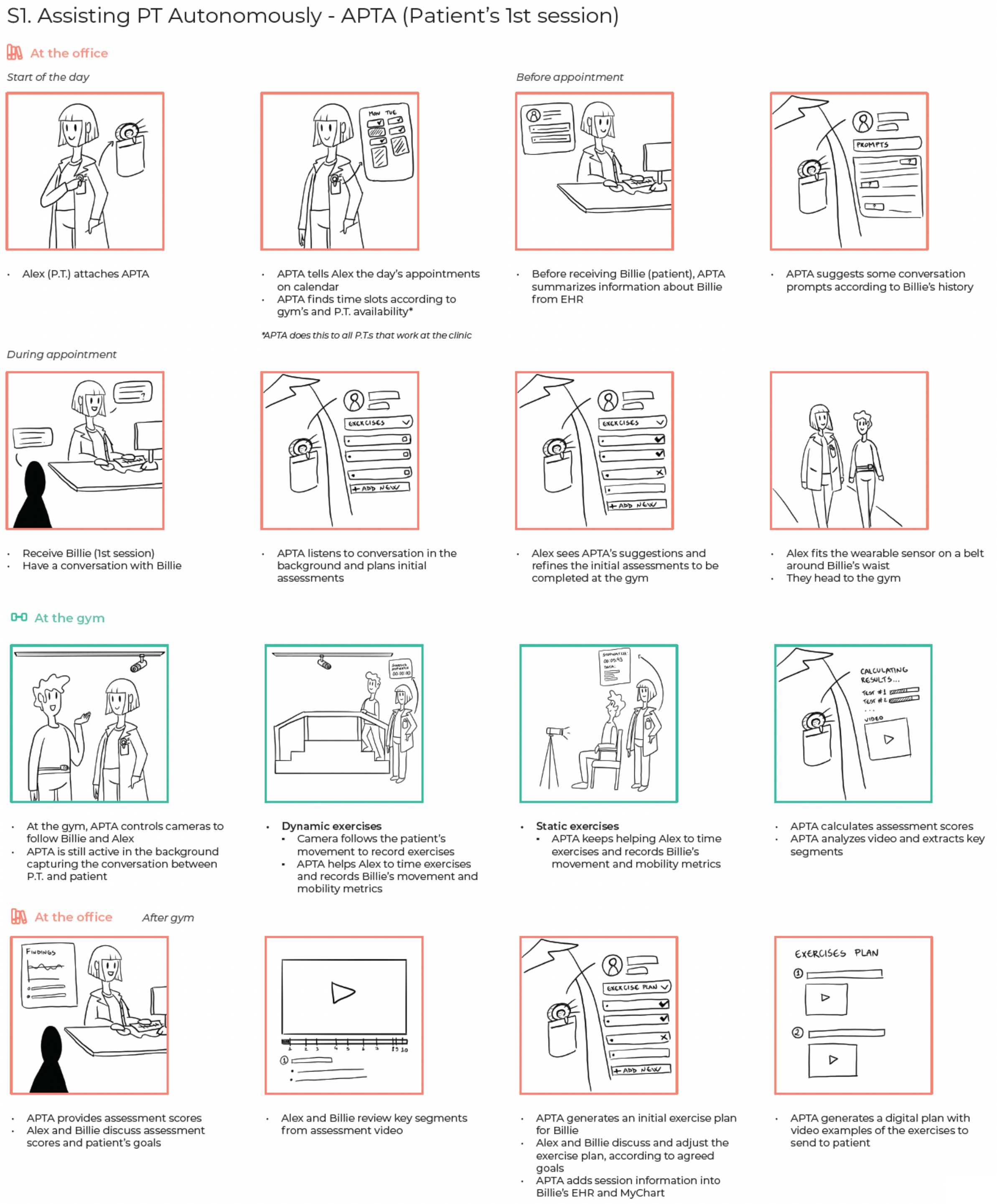
This design kit is intended to support clinicians in exploring how emerging healthcare technologies might better support existing practices, and inspire opportunities for new interactions with patients. The kit includes a pack of inspirational prompt cards, focusing on tools to collect healthcare data and tools to represent healthcare data, and a worksheet to support sharing future healthcare design scenarios. The kit is intended to be used collaboratively by healthcare professionals and designers. To use the kit, clinicians are asked to imagine themselves in near future scenarios and use the cards as inspiration to describe ideal patient-clinician interactions.

The kit includes a single 2-sided worksheet used to structure design activities focused on different scenarios. Each scenario is set 3 years into a fictitious future where technology has developed to the degree that data can be collected, shared, and represented, without the frustrations and IT challenges currently typical.

Figure 7: Digital healthcare design cards representing how to collect data, from the toolkit we created for co-creation activities with PTs.

Figure 8: Digital healthcare design cards for how to represent data, from the toolkit we created for co-creation activities with PTs.

B APPENDIX 2: ADDITIONAL STORYBOARD SCENARIOS


S1. Assisting PT Autonomously - APTA (Patient's 1st session)

At the office

Start of the day

- Alex (P.T.) attaches APTA

- APTA tells Alex the day's appointments on calendar
- APTA finds time slots according to gym's and P.T. availability*

*APTA does this to all P.T.s that work at the clinic

Before appointment

- Before receiving Billie (patient), APTA summarizes information about Billie from EHR

- APTA suggests some conversation prompts according to Billie's history

During appointment

- Receive Billie (1st session)
- Have a conversation with Billie

- APTA listens to conversation in the background and plans initial assessments

- Alex sees APTA's suggestions and refines the initial assessments to be completed at the gym

- Alex fits the wearable sensor on a belt around Billie's waist
- They head to the gym

At the gym

- At the gym, APTA controls cameras to follow Billie and Alex
- APTA is still active in the background capturing the conversation between P.T. and patient

- **Dynamic exercises**
  - Camera follows the patient's movement to record exercises
  - APTA helps Alex to time exercises and records Billie's movement and mobility metrics

- **Static exercises**
  - APTA keeps helping Alex to time exercises and records Billie's movement and mobility metrics

- APTA calculates assessment scores
- APTA analyzes video and extracts key segments

At the office After gym

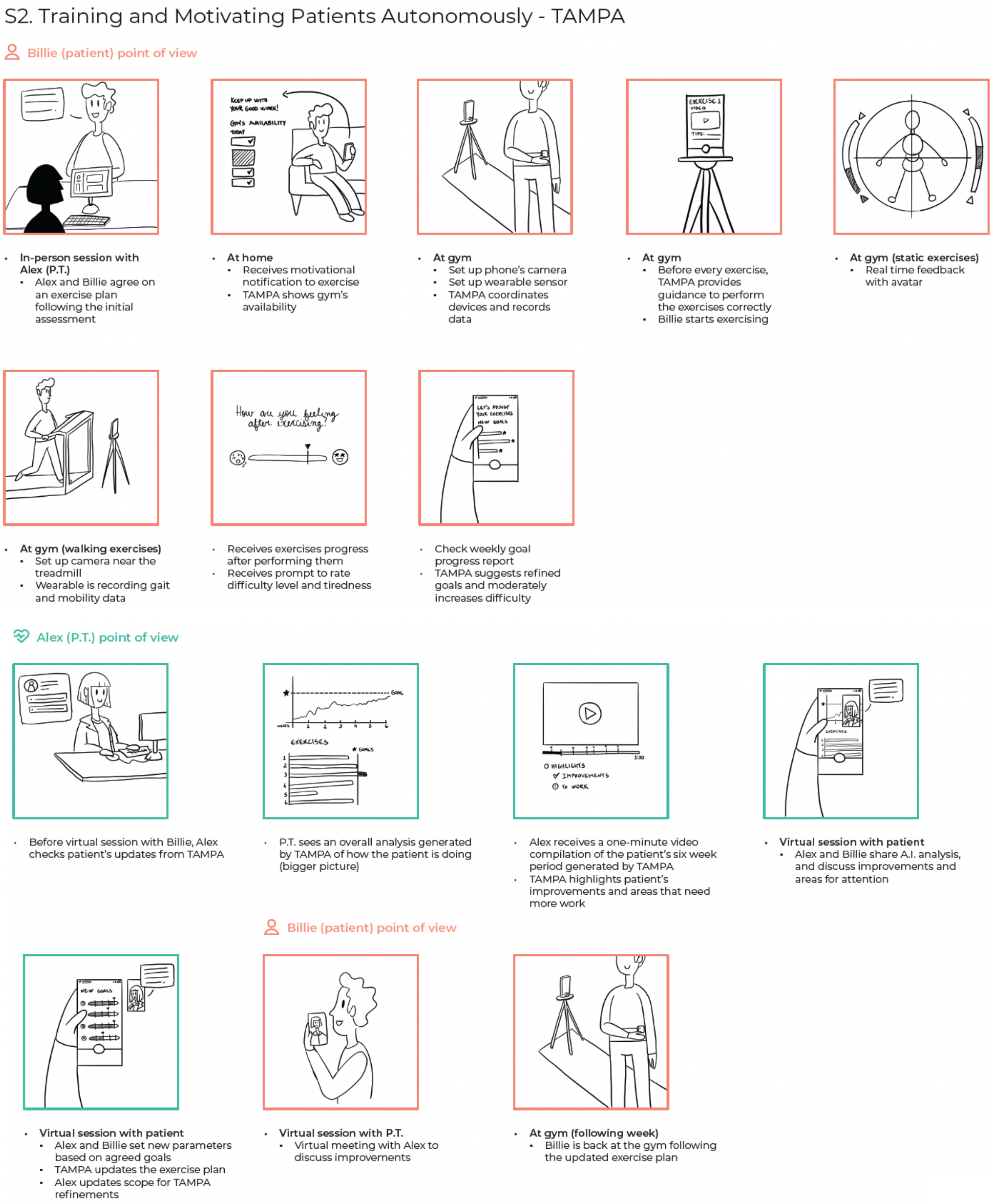
- APTA provides assessment scores
- Alex and Billie discuss assessment scores and patient's goals

- Alex and Billie review key segments from assessment video

- APTA generates an initial exercise plan for Billie
- Alex and Billie discuss and adjust the exercise plan, according to agreed goals
- APTA adds session information into Billie's EHR and MyChart

- APTA generates a digital plan with video examples of the exercises to send to patient

Figure 9: APTA, a system for proactively providing information and managing support tasks during in-clinic sessions, responds to commands in a variety of modalities, including voice for hands-free interaction. APTA also collects and evaluates data from diverse sensors, and summarizes patient-clinician conversations and medical information, and coordinates activities across a group practice or shared gym.



S2. Training and Motivating Patients Autonomously - TAMPA

Billie (patient) point of view

- **In-person session with Alex (P.T.)**
  - Alex and Billie agree on an exercise plan following the initial assessment
- **At home**
  - Receives motivational notification to exercise
  - TAMPA shows gym's availability
- **At gym**
  - Set up phone's camera
  - Set up wearable sensor
  - TAMPA coordinates devices and records data
- **At gym**
  - Before every exercise, TAMPA provides guidance to perform the exercises correctly
  - Billie starts exercising
- **At gym (static exercises)**
  - Real time feedback with avatar
- **At gym (walking exercises)**
  - Set up camera near the treadmill
  - Wearable is recording gait and mobility data
- Receives exercises progress after performing them
- Receives prompt to rate difficulty level and tiredness
- Check weekly goal progress report
- TAMPA suggests refined goals and moderately increases difficulty

Alex (P.T.) point of view

- Before virtual session with Billie, Alex checks patient's updates from TAMPA
- P.T. sees an overall analysis generated by TAMPA of how the patient is doing (bigger picture)
- Alex receives a one-minute video compilation of the patient's six week period generated by TAMPA
- TAMPA highlights patient's improvements and areas that need more work
- **Virtual session with patient**
  - Alex and Billie share A.I. analysis, and discuss improvements and areas for attention

Billie (patient) point of view

- **Virtual session with patient**
  - Alex and Billie set new parameters based on agreed goals
  - TAMPA updates the exercise plan
  - Alex updates scope for TAMPA refinements
- **Virtual session with P.T.**
  - Virtual meeting with Alex to discuss improvements
- **At gym (following week)**
  - Billie is back at the gym following the updated exercise plan

Figure 10: TAMPA, a system for personalized training support and managing exercise programs, maintains an exercise calendar, offers reminders and prompts to motivate exercise, collects and analyzes balance and mobility data, provides real-time feedback to help patients improve performance, fine-tunes programs within parameters set by the physical therapist, and summarizes multiple sessions into key highlights to support patient-clinician discussions and goal adjustment.