Designing for Reflection in Telehealth: Chronic Obstructive Pulmonary Disease Patients as Self-Trackers

Stephanie Githa Nadarajah

Aalborg University

Rendsburggade 14

9000 Aalborg, DK

snadar11@student.aau.dk

Peder Walz Pedersen

Aalborg University

Rendsburggade 14

9000 Aalborg, DK

pwpe08@student.aau.dk

ABSTRACT

Telehealth systems for early detection and treatment of chronic conditions have seen increased use. But the effects on user needs and concerns when healthcare provider continuously monitor and patients provide subjective and objective data over time is poorly understood. Personal Informatics literature informed the analysis of interviews with six Chronic Obstructive Pulmonary Disease (COPD) patients to improve understanding of user needs and concerns in the use of a state of the art telehealth solution. While patients generally felt taken care of, the system in many ways did not meet user needs, e.g. due to difficulties assessing reliable subjective measures and no support on reflection and follow-up action. Interviews, workbooks and design feedback sessions with patients served as the foundation for redesigning the system to support data collection and reflection. Findings from a two week trial involving five COPD patients showed that the system supported one of two types of patients in becoming more informed and aware about their health status, leading to increased empowerment in their everyday life and motivation to set goals and improve condition.

Author Keywords

self-tracking; chronic obstructive pulmonary disease; COPD; personal informatics; data collection; reflection; telehealth;

INTRODUCTION

In response to ageing societies, there is an increasing need for people to take an active role in their own health and well-being. Designing technologies that support self-reflection, awareness and self-management of chronic conditions have particularly been of interest to human-computer interaction researchers for that particular purpose. Chronic Obstructive Pulmonary Disease (COPD) is a progressive lung disease in which the airways are damaged. People with chronic health conditions such as COPD occasionally experience exacerbations that, if not detected and treated early, result in an increased use of healthcare services and a decline in health-related quality of life. Previous studies have shown poor self-management among COPD patients, who do not respond to early warning signs [28, 13, 3]. Patients have difficulties recognizing the onset of an exacerbation and its importance, resulting in delayed recognition and treatment of exacerbations.

Telehealth technologies where patients log and keep track of important health-related variables and symptoms (e.g. self-reported shortness of breath or fluctuations in oxygen saturation measures) are widely used today by healthcare providers to remotely monitor patients and support early detection and initiation of treatment. We refer to the ongoing activity of logging data (objective or subjective) during concrete *episodes* over time as *self-tracking*. Self-tracking is well-understood from literature on the Personal Informatics movement, where people voluntarily, self-initiate and successfully use technology to enhance self-awareness, self-reflection and thereby change behaviours, improve health or other aspects of life. While telehealth literature provides little information on, how patients are supported in their self-tracking efforts, Personal Informatics literature investigates interaction design aspects of self-tracking important for the successful integration of technology in self-tracking interventions.

We synthesized findings from the disparate fields of telehealth and Personal Informatics and used it as an analytical lens for material from semi-structured interviews with COPD patients using a telehealth solution to understand, how we can support patients in their self-tracking efforts (Study 1). While patients generally felt safe, many struggled with subjective data entries that employed references to perceived baselines. The state of the art telehealth system patients used provided little support on follow-up actions and reflection by excluding access to historical data.

We used findings from the literature on Personal Informatics and Study 1 to propose design solutions in an initial prototype, focusing on supporting self-reflection. We conducted individual feedback sessions on the prototype with five COPD patients (Study 2) and found that not all patients were willing to engage in self-reflection and suggested concerns related to the design proposal.

We redesigned and implemented a new prototype for evaluation in the real context during a two week trial involving six new COPD patients (Study 3). The trial informed, how the concrete design decisions regarding entry and interaction with data affected reflection among the patients. While some patients were not interested in reflecting on their self-tracked data as found in Study 2, other patients benefited by becoming more informed and aware about their health status, leading to increased empowerment in their everyday life and feeling motivated to self-improve by setting goals.

BACKGROUND

In recent years, Personal Informatics or Quantified Self (QS) tools have received an increasing interest in the field of human-computer interaction with the introduction of low-cost mobile applications, wearables and advances in sensor technologies. Personal Informatics help people understand themselves through self-tracking of *"personally relevant information for the purpose of self-reflection and gaining self-knowledge"* [14].

Stages of Self-Tracking

Researchers have proposed different models for understanding, how self-trackers concretely use Personal Informatics tools over time. Li et al. proposed the cascading five-staged *Personal Informatics Model* describing, how self-trackers transition between: (1) *Preparation* (determining variables, tools and frequency of tracking), (2) *collection* (logging data), (3) *integration* (preparing data for reflection e.g. by aggregating and analysing data), (4) *reflection* (examining data to gain self-knowledge) and (5) *action* (deciding what to do with said knowledge) [14].

Comparing Personal Informatics and telehealth literature, we found differences in roles of stakeholders illustrated on Figure 1. In telehealth, healthcare providers mandate and predefine what symptoms, how often and with what tool patients should track (*preparation*). Patients collect both objective numerical (e.g. oxygen saturation measures) and subjective binary data (e.g. yes/no answers to whether dyspnea has increased more than usual) (*collection*). The system integrates the data and based on predefined individual normal ranges flags data for follow-ups (*integration*). Trained nurses or physicians review the self-tracked data (*reflection*) in telehealth and if needed contact and advise the patient on potential initiation of treatment (*action*) [28, 21].
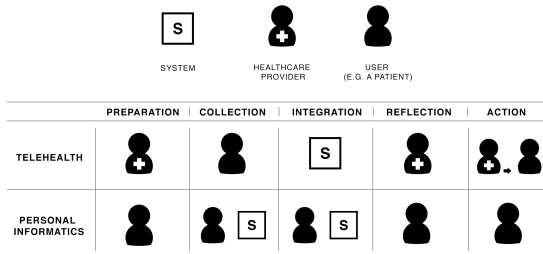

Figure 1. Roles of in Personal Informatics and telehealth

Epstein et al. showed that *collection*, *integration* and *reflection* are ongoing processes that can occur simultaneously and categorized these activities as *tracking and acting*. They further divided the *preparation* stage into *deciding* (to self-track) and *selecting* (a tool to track with) and integrated *lapsing* (e.g. due to an oversight or holidays) and *resuming* into their *Lived Informatics Model* [9, 26].

Rivera-Pelayo et al. discussed Personal Informatics in light of reflective learning theory (or learning by reflection) [24].

Based on the work of Boud et al., they proposed a framework consisting of three dimensions in which technology can be integrated to support reflection: (1) *Tracking*, (2) *triggering* and (3) *recalling and revisiting*. *Tracking* consist of logging data that serves as the basis for the reflective process. Data can both be experiences (e.g. feelings or physiological data), but also outcomes (e.g. gained insight or changes in behaviour). *Triggers* have the purpose of raising awareness and detect discrepancies. These are related to the initiation of reflection based on the logged data. Finally, enrichment (i.e. enriching the tracked data with e.g. context data) and data presentation (i.e. visualising the data) have the purpose of facilitating *recall and revisit* of past experiences.

We use the *Lived Informatics Model* [9] and Rivera-Pelayos framework as an analytical lens to identify user needs and barriers for self-tracking and synthesize findings from literature. We aim to provide an understanding of, how technology has and can be used to support people in their self-tracking efforts in the following section.

Preparation

The *preparation* stage includes people getting motivated to track (e.g. because of a goal they have in mind), which guides the decision on, what data to track and selecting the tool to track with.

Motivation and Goals

Epstein et al., found that people self-tracked for various reasons, but not all people self-tracked with a concrete goal in mind [9]. One example is people, who tracked out of natural curiosity about what their data might reveal about themselves [14, 9]. We refer to this as self-tracking for *life experience*. From the literature, we classified four main drivers for sustained tracking among people with a health-related condition: (1) *documentation* (e.g. to create records for their healthcare providers [1]), (2) *communication* (e.g. to communicate their condition to family members [16]), (3) *self-knowledge and advice* (e.g. to get a sense of the current state of their condition or to get advice from their healthcare-provider [16, 1]) and (4) *self-improvement* (e.g. to change or maintain a behaviour in order to improve well-being or lifestyle [16, 1, 6]).

Barriers to motivation included strong emotional adversity to reflection on data, because it reminded people of negative aspects of their illness [14, 1], tracking the wrong data or not tracking well enough to gain benefits [5], effort [4, 20], reliability and relevance [19, 28, 9] and mismatches between subjective feeling and objective measures [1].

Selection of Data and Tool

Patel et al. found that, sometimes healthcare providers gave little support on what symptoms to track and how to track (e.g. frequency of tracking) [20]. This is an issue in health conditions that involve many symptoms that arise unexpectedly, where patients have to decide on additional symptom tracking [20, 6]. Trackers used tools, such as notebooks, health diaries and specific applications to do additional tracking or sometimes developed tools themselves that were cumbersome and incomplete [20]. This later affected the *integration* and *reflection* stages, where healthcare providers strug-

2

gled with interpreting data sets with additional and not always relevant information [5, 6].

Little is known about what motivates to self-track in telehealth. As previously mentioned (See Figure 1), the decision on what data to track and which tool to use, is in telehealth context decided by a healthcare provider. This potentially has an effect on the acceptance from the user and the efficiency of learning by reflection, as mentioned by Rivera et al. [24].

Collection

The *collection* stage deals with logging data. Self-tracker find logging too many things challenging and leading to tracking fatigue [4, 20]. Logging, however, is a prerequisite for triggering and supporting *reflection*. It is thus of importance to choose as unobtrusive a method as possible that is sufficiently reliable [18]. We previously described self-tracking as logging data (either objective or subjective) during concrete *episodes* over time. We now distinguish between data being either *automatically* logged (e.g. physiological sensors that log heart rate) and *manually* logged (e.g. self-reporting feelings or self-measured value).

Manual Logging

Manual logging requires responsibility and motivation that has to be kept over time from trackers and can therefore be burdensome compared to automatic logging [14, 18]. Some self-trackers found manual logging time consuming and requiring effort [1], hampering incorporation of tracking into daily routines [29, 1]. E.g. when objective data had to be manually logged, it required preparation time, e.g., resting before taking physiological measures, which increased time and effort [29]. Data granularity impeded logging, because trackers had to overthink in order to rate mood on a scale from 0 to 10 [19] and lack of baseline to compare with caused difficulties for trackers with chronic conditions, when self-reporting on severity of symptoms, because they were constantly symptomatic [28]. Another downside of manual logging is that self-trackers sometimes postpone manual logging [16], which induces recall bias that affects data reliability.

Automatic Logging

Automatic logging shifts the effort from the trackers to the technology. It allows for configuration of frequency and precision, but also implies challenges in terms of filtering and aggregating the large amounts [18]. One of the main downsides of automatic logging is that it might reduce awareness and self-reflection [4, 15].

Integration

The *integration* stage can be more or less apparent to the tracker depending on whether integration is automatically done by the system, requiring less effort from the self-tracker [14]. Previous studies show that self-trackers sometimes postponed data exploration, when integration did not happen automatically, since it involved tedious tasks, such as cleaning up data, formatting and running statistical tests [4, 5, 14].

Systems that require manual integration expect the user to be able to analyse the data and ascertain the best way of creating a representation. Whooley et al. found that this is of interest to curiosity-driven self-trackers that want to integrate data manually and explore the novel insights that data can offer them. In contrast, self-trackers with a goal, knew what they were looking for in the data and strived at using automatic integration systems, allowing them to concentrate on reflection. The manual integration process is an iterative process of moving back and forth between representation creation and reflection [30].

Reflection

In the *reflection* stage, self-trackers reflect on the collected and integrated data by looking, exploring or interacting with visualisations of it [14]. While literature points at many different definitions of reflection, we found little about how reflection can be evaluated. Fleck & Fitzpatrick identified five different levels of reflection (R0-R4) that indicate what types of activities and behaviours can be associated with reflection [10]. Levels consist of (R0) describing or stating without being reflective (*description*), (R1) describing with explanations in a reportive or descriptive way (*reflective description*), (R2) seeing things from different perspectives and trying to identify relationships (*dialogic reflection*), (R3) changing original point of view due to gained knowledge (*transformative reflection*) and (R4) seeing the wider perspective beyond the immediate context (*critical reflection*). While higher level indicates being more reflective, lower levels are prerequisites for becoming more reflective.

Conditions for Reflection

One of the condition for reflection is creating and allowing for time to reflect, according to Fleck & Fitzpatrick [10]. Li et al. distinguished between short-term reflection, where the self-tracker reflects immediately after logging the data and long-term reflection that might occur several days or weeks after [14]. While short-term reflection makes the tracker aware of the current status, long-term reflection allows for higher levels of reflection (at least R2), since the tracker can compare logged data between different times and explore trends and patterns.

Several authors mention that the one reflecting should be open-minded and willing to reflect [2, 25]. As mentioned by Atkins et al., in some literature there is an implicit assumption that *skills* (such as critical analysis and evaluation) are necessary to engage in reflection [2, 25]. Fleck & Fitzpatrick mention that reflection skills can be developed with time and with the right support [10].

The reflective process needs a trigger. People often need a reason (e.g. a purpose) or at least an encouragement to reflect [10, 17]. In psychology, Festingers cognitive dissonance theory describes, how a mismatch (psychological discomfort or dissonance) between an individual's attitude and behaviour can lead to rethinking ones attitude and behaviour [24].

Rivera et al. mention that dissonance triggers reflection and can be actively triggered (system explicitly tries to catch the users attention by highlighting a certain mismatch) or passively triggered (system only presenting the data and relying

on the user to detect something that starts a reflective process). Dissonance may occur due to comparison between current level and a recommended level or goal, but some might prefer such comparison in response to oneself. For example, Ancker et al. found that some self-trackers prefered to interpret their data in light of their own personal and medical history and/or symptoms, rather than striving for provider-recommended normal ranges [1]. In the following we describe different ways to trigger and support reflection identified in our literature review.

Reflective Questions or Prompting
One way of supporting reflection is through the use of reflective questions or prompts. Simply asking people to provide justification or explanation for e.g. events or actions can support *reflective description* (R1) [10]. The presence of another person can encourage reflection, especially in a dialogue among two uneven partners (i.e. two people not sharing the same understanding or experience), where one takes the role of asking questions.

Systems can also take this leading role of asking questions, but opposed to the previously mentioned example, where reflection among two people can be dialogue driven, systems often only pose an initial reflective question [17]. An intelligent system could further support reflection through follow-up questions. Another ways to foster reflection is by prompting questions triggered by automatically logged context data [10]. Presenting data that is not usually visible encourages people to see things from another perspective and can potentially lead to looking for relationships and patterns (R2) according to Fleck & Fitzpatrick.

Visualisations
Despite simply looking at data is not considered reflective according to Fleck & Fitzpatricks levels of reflection, creating representation of data is considered a prerequisite to support higher reflective levels [10]. Visualisations of data can help people in exploring their information and gaining insights [14, 4].

When designing visualisations, it is important to consider under which conditions reflection is to take place (e.g. time and effort expected from the user, skills and purpose for reflection) [7, 18]. Li et al. identified six types of questions people ask about their self-tracked data [15]. These are, getting to know (1) status (what is current status?), (2) history (what has status been in the past?), (3) goals (what goal is appropriate to pursue?), (4) discrepancies (how does current status compare to goal?), (5) context (what affects current status?), (6) factors (how are different variables related?). Depending on the conditions, supporting both simple (e.g. status charts) and detailed visualisations (e.g. of time series) can be important [18, 7].

Often people want to obtain answers to their question (e.g. status) without spending too much time or effort, which can be done on a simplified dashboard representation that allows for a quick overview [7]. Müller et al. found that people used status charts to quickly get an overview of their data and used it as a starting point for exploration [18]. Compari-

son charts were requested in their study to benchmark against other people in the sense-making process and to assess success. However, as previously mentioned, some prefer that such benchmarking occurs in response to oneself [1].

Previous literature shows that visualisation of time series data can support revisiting and analysing past experiences (history) and trigger storytelling about experiences behind data [24, 18]. It can foster reflection on global trends e.g. on upward and downward trends or deviations from a historical normal (suggesting a problem) [24]. Cuttone et al. mentioned that human behaviour is often characterized by periodic patterns, but that time series graphs do not facilitate exploring such patterns [7]. They instead proposed using calendar heatmap representations using different color shades to indicate variable values. Visualisations of multiple time series can support reflection on how multiple variables are related or how multiple variables change over time [7].

Visualisations of time series can further be combined with discrete events [27] to support reflective description (R1). Previous studies showed barriers for reflection involved that tools did not always support either simple visualisations of data or more complex features (e.g. filtering data to focus on a subset of data, zooming out to get an overview or comparing multiple variables) [15, 16].

External activities
While reflection is an internal process, it can occur when trying to externalize thoughts and feelings e.g. in diaries or during reflective writing [17]. These activities are often descriptive or emotional (R1). Recording reflection outcomes for later revisiting and reflection on gained insights has been proposed as another way to support reflection [12, 18]. Isaac et al. found that allowing for writing down thoughts helped people explore and understand their feelings [12].

Action
People decide what *action* to take based on the findings from the *reflection* stage. Epstein et al. found that trackers who had other motivations than behaviour change, e.g. self-understanding, did not reach this stage [9]. Trackers who tracked for self-improvement sometimes lacked the knowledge necessary to identify the appropriate actions to take, such that they could regulate their progress towards their behaviour change goal. This happened either because they collected irrelevant data [4, 5] (e.g. food and symptoms, instead of ingredients that triggered the symptoms) or because they needed actionable (expert) advice [29, 14, 19]. Several studies found that most Personal Informatics systems do not provide actionable advice [5, 14, 29].

RESEARCH AREA
In summary, previous literature shows user needs and concerns during tracking activities, but little is known about these aspects, when tracking is monitored by a healthcare provider as in telehealth. It is unclear how concrete design decisions regarding entry and interaction with data affect reflection. In the following studies, we explore user needs and concerns when self-tracking in telehealth and how we can design to support users' self-tracking needs.

STUDY 1 - EXPLORING USER NEEDS IN TELEHEALTH

In the following study we explore self-tracking needs and concerns in telehealth context using COPD patients as our case.

Participants and Method

Six COPD patients (P1-P6), two male patients (P1, P2) between 64 and 65 years (M: 64.5) and four female patients between 54 and 74 years (M: 66.8), participated in the study. We established contact with patients through Silkeborg Regional Hospital. All patients had either severe or very severe COPD and multiple other health-related conditions (diabetes, heart disease, pulmonary oedema, asthma, bronchitis, osteoporosis and sleep apnea). Three of the patients used supplemental oxygen (P3, P5, P6). All of them lived in their own homes with a spouse, except P3 who lived alone. P6's spouse (P6S) was her spokesman, as she had a speech disorder and took responsibility for her.

We conducted audio-recorded semi-structured interviews in patients' own homes using an interview guide as the framework for discussion. The participants demonstrated to an interviewer and a researcher who took field notes, how they used a telehealth system AmbuFlex (AF). Participants had used AF in between three months and two years. Three of the participants had previously used Tunstall HealthCare (THC) solution (for six months).

AF is web-based and can be accessed by mobile or desktop. THC consisted of a monitoring box installed in the patient's home. In both systems, patients submitted objective data (oxygen saturation, lung function (only in THC), pulse and weight) and subjective data (binary answers to whether dyspnea, cough and sputum color had been higher than usual) three times a week (Monday, Wednesday and Friday). AF provided confirmation of submission but no option for reviewing previous data. Patients received follow-up calls when healthcare providers monitoring the data needed additional information (e.g. in order to discuss deviations and/or advise patients' to initiate medication) or validation of measures.

We went through the field notes and audio-recordings after the interviews. In the following we present, how patients engaged in the different stages of self-tracking using AF.

Results

All patients remembered to take their measurements consistently and routinely in the morning themselves.

Preparation:

The majority of participants (P1-P5) found the sense of security from healthcare providers monitoring their data motivating. *"(...) it gives you a huge sense of security that you are not gonna lay at home ill"* (P2). P4 felt obligated to take measurements due to the presence of a healthcare provider.

Two patients tracked additional data on paper (P5, P6S). P5 used the data as documentation, e.g. when being admitted to the hospital to discuss it with healthcare providers. P6S' spouse mentioned curiosity, self-satisfaction and sense of agency as motivations for tracking on paper.

Collection:

Patients found data collection easy, not requiring expert computer skills, and not taking too much effort or time. P5 stressed the importance of fast collection, *"it must not take ten or fifteen minutes to do it everyday (...)"*. *"This [AF] is really simple (...) it is so simple you can add some more to it"* (P6S). Patient mentioned currently spending between 1-2 minutes and 15-20 minutes on AF. Several patients found answering subjective questions difficult when it required comparisons with the 'usual' baseline, (e.g. *"Are you coughing more than usual?"*). *"What is usual? Isn't that also how I felt yesterday? Otherwise, I have misunderstood the question"* (P4). Patients needed higher than binary granularity to answer, *"When they ask if you have more dyspnea than usual, then we say yes .. but how much is it? They [healthcare providers] cannot see"* (P6S). Some patients underreported baseline deviations and only answered "yes" in large or extreme deviations, *"if it's just a little different, I do not mention it"* (P2), *"I would have to be coughing a lot and feel very ill, if I answer yes to that question (...)"* (P4). P2 asked for a scale instead *"(..) why don't they make a scale instead for example from 1 to 5 or 1 to 10? One day I could perhaps say it's 5, the next day 6 and the day after I can go back to 5"*. P5 used the comment box to make small deviations go on record, *"(...) to me it is important that we take every small nuance"*. When to collect, was a concern for P4 whose oxygen saturation measure and pulse depended on her level of activity. She wondered why the system did not take into account external factors related to her condition, *"Do you feel more breathless today? But it does not say anything about the fog outside"* (P4).

Reflection:

Several patients mentioned that an exacerbation comes within a few hours or even minutes, and that they were not able to recognize an onset by using AF. P1, P2, P3 and P5 measured oxygen saturation several times a day to verify their subjective feeling of well-being (P2, P5) or lack thereof. None of the patients felt they learned anything about their disease using AF. *I can feel it [an exacerbation], even if I did not have the monitoring device"* (P3). Patients did not express any concerns waiting for a (potential) call. Most of them had identified "the hours" of the reviewers at the hospital and several a mental model of when a call would ensue. *"They usually do it [review the data] before noon"* (P3). *"I already know when there is going to be a call (..) when the oxygen saturation is too low, the pulse is high and your measures fluctuate, they react"* (P2).

P6S found benefits in tracking data on paper, allowing for understanding his wife's baseline and whether she was deviating from it and getting worse. *"You can see how stable it is .. (..) Let's say she loses weight then I become alert that something is wrong"* (P6S). Patients had not been informed by their healthcare providers about their "normal area" (recommended level) and the AF interface did not communicate it either. Half the patients wanted to know these in numbers. Some of the patients had identified their own "normal area" of oxygen saturation that mapped to not feeling well (usually below 90).

Action:

All patients had received education in self-management of their condition (e.g. breathing techniques), but not all patients gained the same benefits from it and needed actionable advice. To that end some patients (P2, P5) added questions to the comment box. P2 acted on the basis of his oxygen saturation measures, "*when it [oxygen saturation measure] is lower than 93, you do not feel fine (...) then I walk a little slower and take it a bit more easy*". P2 was interested in knowing additional methods to increase oxygen saturation. P6S wanted recommendations on duration for supplemental oxygen use based on her oxygen saturation measures and information about variables e.g. the weather, that could influence her symptoms. P3 and P4 used their oxygen saturation measures to adjust their supplemental oxygen. However, P3 preferred not to initiate treatment including drugs before consulting a healthcare provider, unless in extreme cases of symptoms or unavailability of staff. "*I might to do it [initiate medication treatment] if it [sputum color] was very green, if it was a Tuesday [a day not monitored by healthcare providers], otherwise I wouldn't*".

Discussion

Our patients were highly motivated to track potentially due to the active role of the healthcare provider that provided them with a sense of security not present in previous studies [14, 1, 5]. One patient relied on the monitoring to such an extent that she sometimes delayed treatment, waiting for confirmation from the healthcare provider. None of the patients described the tracking activity requiring too much effort or time. However, effort seemed to be an aspect that should be considered when designing telehealth systems, as some patients both expressed willingness to spend more time than AF required (approximately two-three minutes), but not wanting to spend more than ten to fifteen minutes.

Based on our findings and literature review, we revised the *Lived Informatics Model* [9]. We broke down a data collection episode into *pre-collection* (deciding on whether to log or skip), *acquisition* (ready required artifacts), *calibration period* (satisfy guidelines for tracking), *measure* (taking measure using artifact), *entry* (entering read off measure from artifact or providing scale based ratings, absolute or relative to a baseline, or qualitative comments) and *submission* (submitting data). The patients reflected during multiple stages of collection before entering both subjective and objective measures.

AF did not meet the needs of users in terms of (1) scope, (2) reliability, (3) validity, (4) actionable advice. AF's scope focused only on submitting variables directly related to the condition at time of entry, but two patients logged additional data on paper (c.f. [20, 6]). Having access to their previous data made patients feel in control. In terms of reliability, some patients felt unsure whether they were collecting data under the right conditions. Subjective questions with a baseline comparison proved difficult due to: No access to the baseline and the low granularity of the answer options. Patients had insufficient access to their usual subjective feelings and tried to remember previous events to establish their 'usual' baseline

(c.f. [28]) and AF provided no access to historical data. Even if AF provided access to previous data this might prove difficult due to the low granularity. The binary answer options resulted in reduced validity of data by underreporting significant increases from the baseline. One patient specifically asked for rating on a scale instead, which requires more cognitive effort and time [19]. Due to the absence of data access, patients did not interact with the data they had collected and expressed not having learnt anything from telehealth. Several patients mentioned not being able to recognize onset of an exacerbation from AF use, suggesting that the system poorly supported reflection.

Patients did not know their provider-recommended "normal area" and therefore used their own identified "normal area" in management of their condition using the pulse oximeter. Some patients wanted to know the provider-recommended "normal area" to become more empowered, while others were not interested. One reason for that could be that patients get reminded about the negative aspects of their health when reviewing data or that they rely more on their subjective feeling than on quantities, as in [1].

Several patients were interested in actionable advice from the system as in [5, 14]. Apart from during subjective data entry, patients needed two types of support, (1) confirmation from healthcare providers to act (e.g. initiation of medication) and (2) actionable advice on self-management strategies (e.g. coping with breathlessness). We believe that one of the barriers to action was the lack of support for reflection during entry and review of data - a prerequisite to action according to Li et al. [14].

Conclusion

While having a healthcare provider monitoring data motivated sustained tracking, the telehealth system patients currently used did not sufficiently meet user needs for tracking. Patients expressed difficulties rating their symptoms relative to their usual baseline and uncertainty in terms of which conditions to measure in, resulting in reduced reliability of the data. A lack of access to their historical data hindered patients in entering reliable data, reflecting and taking actions.

STUDY 2 - DESIGNING A COPD TELEHEALTH SYSTEM

In the following we describe how we redesigned the system to support the COPD patients in collecting and reflecting on self-tracked data and evaluated the re-design.

Methodological Considerations

Initially, we planned to conduct co-designing activities with the COPD patients to meet their user needs. However, Das et al. found that co-design sessions with COPD patients using generative tools and techniques (e.g. post-it notes and sketching activities) is resource demanding for patients due to their health condition [8]. Some patients were not even able to participate in other activities than keeping a conversation. In Study 1, we similarly experienced that the COPD patients were physically limited (e.g. in terms of moving from one place to another) and within an hour of interview experienced breathing difficulties several times, demanding a slow pace

and long pauses. Drawing from Das et al.'s experiences and our own experiences from Study 1, we decided not to conduct co-designing activities as initially planned.

We made a conscious plan to conduct individual sessions, where patients were only expected to either talk or interact with a prototype. To further reduce time and effort required by patients in the sessions, patients completed workbook assignments beforehand, allowing them to prepare for the discussion on, how our proposals foster reflection during collection and review of data. In the following, we describe the workbook and prototype design.

Workbook Design
Workbooks assignments aimed at getting insight into, how the system can support patients in collecting and reflecting on data. Study 1 showed that patients were unsure, whether they were collecting data under the right conditions. We asked them to annotate, if they had reflected on context-relevant variables provided in the workbook to get further insight into, how patients reflect on conditions relevant for the validity and reliability of their measures. For dyspnea measures, context-relevant variables could be e.g. weather, mood, smoking or physical activity. Patients commented on it as free text annotation one day and mark it in checkboxes another day (e.g. mark with X if medicine has affected saturation measure).

Figure 2. Workbooks with assignments

In Study 1, patients expressed difficulties rating symptoms subjectively, due to lack of baseline understanding and low granularity of answer options. Patients had to reflect on a time series graph of previous measures (showing baseline) while entering current measure in one assignment. In another assignment, patients assessed dyspnea on different scales: (1) Relative to usual baseline (binary), (2) absolute (five-point scale inspired by the questionnaire-items in EXACT PRO [11]) and (3) using visuals (Dalhousie Pictorial Scale [22]).

Finally, patient had to reflect on a time series graph of previous oxygen saturation measures (long-term reflection). A recommended level (also referred to as goal or "normal area") was marked on the graph to trigger reflection on a potential mismatch in accordance with Festinger's cognitive dissonance theory [24].

Prototype Design
Improving Reliability and Validity of Measures
The collection page had features similar to the workbook assignments (See Figure 3). A time series graph visualising patients' previous measures supported patients in remembering previous events while entering data in the prototype.

Symptom rating options were scale-based instead of binary as patients were used to, inspired by questionnaire-items in EXACT-PRO [11]. We further introduced the option to toggle context-relevant variables for each measure in the prototype to provide enrichment [24], support reflection while collecting and later review (description follow in next section).

Supporting Short-term and Long-term Reflection
From literature, we found that supporting both short-term and long-term reflection can be important, depending on the conditions in which the user might self-track [14, 18].

To support short-term reflection, we provided patients with a dashboard view after collection (See Figure 3). The dashboard view aimed at allowing patients to reflect on current status immediately after collection and increase awareness on their current status [7, 18]. We included reflective questions to trigger reflective description (R1) or higher levels of reflection on this view and to encourage users to further explore their data [10, 18]. E.g. *"Why are you coughing more than last time you measured?"*. Gauges for each measure showed current measure in relation to the recommended level and arrows indicated change from previous measure (day-to-day variations). The system used color indications (red, yellow and green), where red or yellow highlighted a potential mismatch and thereby aimed at actively triggering reflection in accordance with Festinger's cognitive dissonance theory [24].

To support long-term reflection, we designed four different visualisations (See Figure 3). (V1) Time series graphs with option to compare two measures, (V2) Time series graphs stacked vertically, (V3) Calendar heatmap and (V4) Area graphs with option to compare two measures. V1, V2 and V4 were proposed to foster reflection on upwards and downwards trends or symptoms deviations to increase awareness on a worsening in condition [24]. V1 and V5 differed in that they allowed for comparison of multiple measures that could trigger reflection on, how measures were related or changed over time [7], affording dialogic reflection (R2) [10]. Recommended levels were highlighted on all four visualisations to increase awareness on discrepancies [14] and trigger reflection. V3 was proposed as an alternative to time series graphs as mentioned by Cuttone et al. to support reflection on periodic patterns using color shades to indicate deviations from recommended level [7, 14].

To further support long-term reflection, we provided the option to toggle on discrete events on the visualisations (context-relevant variables) [27]. This feature was included to support reflective description (R1) and dialogic reflection (R2) by allowing for exploration of relationships between context-relevant variables and measures otherwise invisible [10].

Participants and Method
We asked the same patients as in Study 1 to participate in feedback sessions on the redesigned system. All patients and three spouses (P2S, P5S and P6S) participated, except P4 who was hospitalized. Feedback sessions were held in the
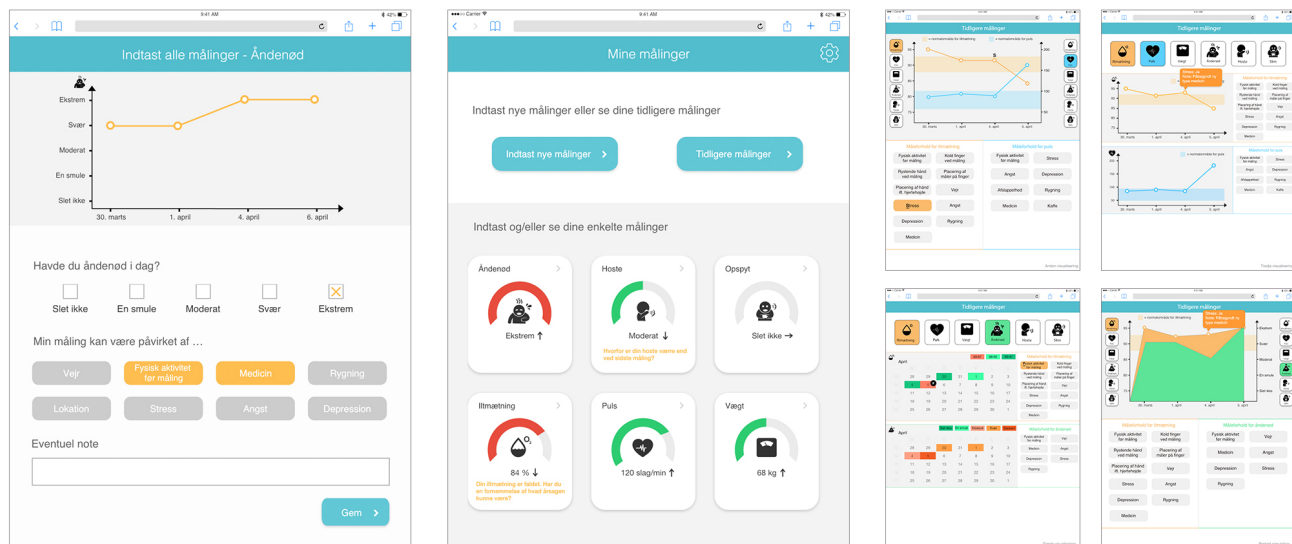
Figure 3. Screens from the prototype: Collection page (left), Dashboard (middle), Four visualizations of history data (right): V1 (top left), V2 (top right), V3 (bottom left) and V4 (bottom right)

patient's home and audio-recorded. Sessions lasted approximately one hour.

Patients completed workbook assignments three days a week (Monday, Wednesday and Friday) delivered one week prior to the feedback sessions. Feedback session were divided into a workbook session and a prototype session, where two researchers were present. One researcher took the facilitation role and asked patients follow-up questions emerged from Study 1, while the other researcher went through the workbooks and prepared for the workbook session. After the follow-up, the researchers shifted roles. The researcher responsible for the workbook interviewed patients and discussed assignments, while the other researcher prepared the prototype. The prototype preparation involved adjusting the prototype to use patient's own values for visualisations (collected in workbooks) to encourage reflection on patients' own experiences rather than fictive values. In the prototype session, patients went through each screen in the prototype and completed usability tasks and discussed understanding and thoughts on the prototype. The researcher encouraged both patients and their spouses to think-aloud and provide their points of view.

Results

We identified two types of patients in this study: Active and Passive. Passive patients (P1 and P3) took the role as data providers and did not see any benefits in neither short-term or long-term reflection provided by the system. Active patients (P2 and P5) were more open-minded towards engaging in the discussion of reflecting on self-tracked data.

Reliability and Validity of Measures

Not all patients were consistent in terms of *when* and *how* they took measures. Some patients had specific routines (e.g.

always before breakfast), while others were not consistent (e.g. P1 only weighed himself every three weeks).

Passive patients and P6S lacked knowledge on relevance of context and how some of the provided context variables influence their measures. E.g. P1 and P6S questioned whether and why a cold finger when taking oxygen saturation measures had an influence. Active patients ensured taking measures under comparable conditions. Patients mentioned additional context variables relevant for their measures (e.g. mood, talk, supplemental oxygen). P2 and P6S specifically asked for guidelines on taking measures. *"For some measures, an explanation would be good. For example do not take measure if this and that"* (P6S).

All patients preferred higher granularity options when rating symptoms more than binary. P5 and P6 mentioned that they preferred higher granularity, because it made it possible to show degrees. *"How much is a no? If we say yes or no to the hospital, they still do not know what we are thinking .. then they'll call us and we'll have to explain it is severe"* (P6S).

Asking to rate symptoms without a baseline comparison (*"Did you feel breathless today?"*) and having five answering options (not at all, slightly, moderately, severely, extremely) or visuals as options caused difficulties for some patients, because they experienced several of the options during the day. *"I have been through all of the provided options that day. How do you want me to answer that?"* (P5). Patients further had different perceptions of what usual is, when asked to rate relative to usual baseline. *"(..) I base that on when I'm at my best"* (P3), *"usual is when it is an ordinary day"* (P5) and *"If she is not more breathless than yesterday, then we'll just submit a no"* (P6S).

8

Patients presumed that it would not be of benefit to them to have a time series graph showing their baseline while collecting. *"I do not think it has any effect to see a graph"* (P3). P2 mentioned, *"I shouldn't answer based on what I answered last time. I should answer what it is now and here"*.

Short-term and Long-term Reflection

In the prototype session, some patients were not willing to reflect on visualisations of current status, reflective questions or history data. E.g. *"I do not care what my status is.. I just submit data.. do not walk around and think everyday .. I know how I am feeling"* (P3). P5 expressed interest in current status and what action should be taken, *"I'm more concrete. Where am I right now and what can I do about it?"* (P5). The dashboard view helped some patients in getting an quick overview of their health status (e.g. P5). P2 mentioned that the arrows helped him *"quickly see if it [a measure] is going up or down"*. Reflective questions were not noticed by the patients. In passive patients reflective questions did not trigger any reflection (R0), while active patients proposed lower level explanations as answers (R1), when we prompted them to answer the question in the system. E.g. to the question *"Why are you coughing more than last time you measured?"*, P5 answered: *"Right now it is likely because I talk too much"*.

None of the patients expressed interest or benefits in having access to previous measures prior to the feedback session, e.g. *"I do not need it [access to history data]"* (P3). Several patients relied on the healthcare monitoring *"I have a nurse who is good at keeping an eye on me"* (P5), *"I do not need it [access to history data]. If it [measure] is too low, they call and ask me why"* (P1). P2 similarly expressed that he did not have the necessary knowledge to find it useful *"I do now know what they use it for, the scales they use and the language.. I do not understand it. I count on them reacting if there is anything"* (P2).

Providing patients with a recommended level that they could compare their measures against, provoked negative feelings among some: *"I prefer not to be told in the morning that I'm gonna get an awful day"* (P5), *"It's ok if it's just a single measure, but if it is constant, I would start thinking.. It's going fast now"* (P2). While P5 mentioned that she would ignore the recommended level, because it did not match with her own goal, *"I thought, you can forget it [about provider-recommended level]. I'll just do what I usually do"*, P2 suggested that he would strive to keep his measures within the recommended level, *"then it's not that bad if I keep it above that [lower threshold]"*. P6S indicated that if his wife's oxygen saturation was above normal area, he would start wonder whether the oxygen supply was set too high and initiate action. *"If it starts to go over here [below normal area], we have to do something"* (P6S).

Despite none of the patients initially expressed benefits in having access to history data, two patients (P2 and P5) changed their attitude after the prototype session. *"This gives more information about me (...) it's nice to be able to go back.. Is it better than 14 days ago?"* (P2). One patient mentioned needing a purpose and time to reflect for gaining any benefits from history data, in accordance with Fleck & Fitz-

patrick [10]: *"There might be days where I sit with it and have an idea about what I'm looking for, which might trigger some thoughts"* (P5). Others were more reluctant on reflecting on history data. P3 did not see any benefits and found it troublesome, while P1 did not think it was his job to look at history data, *"this is only for people who has to sit and analyse the numbers"* (P1).

V1 afforded finding relations between measures e.g. by comparing, *"you can have them (measures) together and see how they affect one another"* (P2). V3 was more attractive to others, who found it more concrete and provided a quick overview, *"it (V3) is the one I get the quickest.. If you are in doubt what the colors mean, you can see them down there"* (P5). PS6 thought that he would start with V2 and then use V1 after becoming more advanced.

Assignments in the workbook and prototype elicited only limited findings on patients' reflective thoughts on the design proposals. While we did use patients' own values rather than fictive values to discuss the workbook and prototype assignments in terms of long-term reflection, we used fictive values on the dashboard, which could have been a barrier for reflection in some patients. Additionally, patients might have found it difficult to reflect, because (1) the interview period was time-limited and (2) a reason (e.g. a worsening in health status) was needed to trigger reflection. Taking into account the findings from this study, we decided to implement the system with a few adjustments and evaluate it on other COPD patients in a two weeks trial.

STUDY 3 - EVALUATING REFLECTION DURING USE
In this study, we explored how interacting with the redesigned telehealth system in a real context affected reflection among COPD patients and their activities in self-managing their condition.

Prototype Re-Design
We made adjustments to the prototype based on findings from Study 2 and developed a web application. An introductory dialogue box provided basic information to increase awareness on the importance of measuring under comparable conditions, context-relevant variables and sudden changes in symptoms that can be indication of an exacerbation.

All reflective questions targeted patients' overall health status, instead of being specific to measures as in Study 2, making it possible to highlight reflective questions in the interface more than before. E.g. *"Have you previously been able to improve your measures? How?"*. Some targeted features of the systems to increase awareness on symptom changes, e.g. *"You have multiple measures showing red/yellow. Is there any improvement in your latest measures? (Look at the arrows on each measure)"*.

A settings option in the system allowed patients to adjust their "normal areas" (recommended levels) to their own preference. We chose V1 for long-term reflection based on the idea that it could be of interest for patients interested in seeing relationships between measures and potentially support awareness on deviations [24, 7].

Methodological and Ethical Considerations

We had a number of ethical considerations on the design and methodology prior to the study. From the beginning, we explicitly told all patients that the purpose of the study was to investigate, how patients reflect on symptom changes and not on providing any kinds of support on action. All patients signed a consent form agreeing that they understood and accepted that their data was not monitored by healthcare professionals as in the telehealth intervention they previously or currently used. We further provided all patients with a hotline to one of the researchers, which they could call if they had any questions related to the system.

We chose to conduct semi-structured interviews as our main method for data collection, focusing on the same methodological considerations as in Study 2. We included logging in the system and diaries as unobtrusive methods for collecting data about, how patients interacted with the system and reflected on their disease during the trial period. We made diary writing optional to decrease effort and ensured that patients prior to the study in the signed consent form accepting that we logged their anonymised data stored on a secured server.

Participants

Five COPD patients (P1-P5), two male patients (P2, P4) and three female patients between 67 and 80 years (M: 73.6) completed the study (See Table 1). Seven COPD patients initially enrolled, but one male patient dropped out due to a bad day and because the study involved a tablet, while another female patient was hospitalized during the intervention. Patients had lived with their diagnosis between 7 and 25 years (M: 12) at the time. All patients had either moderate, severe or very severe COPD. Two of the female patients (P3, P5) used supplemental oxygen. P3-P5 had multiple other health-related conditions (diabetes, osteoporosis and fibromyalgia). P4 reported color blindness, but ability to distinguish between the colors used in the system, while P5 reported cataract affecting her vision.

P1 and P5 lived alone, while all the other patients lived with at least a spouse. None of the patients were active on the labour market at the time, but had been holding diverse professions (e.g. teacher, mechanic, cleaning assistant). P2 was the only patient who did not use technology (e.g. computer, tablet or similar) at all. His wife (P2S) was in charge of helping him and also participated in the follow-up. Based on self-assessment, all the other patients used technology either on a daily basis (P1, P4-P5) or 1-6 times weekly (P3) e.g. to check bank accounts, social media, entertainment and similar. P5 had no prior experience with tablets. P1 and P4 had used AF for three months and continued use during our intervention, P2 and P5 previously used THC (for approximately 6 months). P3 had been in a telehealth intervention, where she participated in the control group.

Procedure and Analysis

We recruited participants through contact with Silkeborg Regional Hospital. A nurse at the hospital called potential candidates from their database and asked for permission to pass contact information to us. Our only inclusion criteria was that

patients had prior telehealth experience. We later found that P3 had only been in the standard care group of a previous telehealth intervention. Our procedure included three steps:

Preliminary Data Collection: After initial consent over phone, we sent out a letter to all patients with a study information sheet and a template, along with pulse oximeter and weight scale if patients did not own it already. For each patient to see their own data from the previous week during the trial, patients had to fill out the template with measures of oxygen saturation, pulse, weight and self-reported dyspnea, cough and phlegm for two days prior to the trial.

Trial Period: We scheduled initial meetings with patients ensuring that they all had a complete TeleKOL kit (See Figure 4) consisting of a pulse oximeter, weight, diary and iPad with internet access and signed a consent form. A facilitator asked patients height information needed in the system to compute BMI (as min and max weight in the system was defined using recommended BMI for COPD patients) and educated patients in the use of the system such as: Opening the application, submitting data, accessing previous measures and settings. We encouraged patients to use the system at least three times a week along with the diary writing. Patients who already used AF were asked to use the system on days where they did not use AF. Debrief with patients occurred 14 days after the trial start.
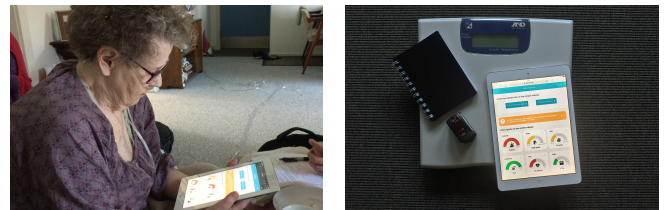

Figure 4. Patient using TeleKOL (left) and TeleKOL kit (right)

Debriefs: Two researchers conducted audio-recorded semi-structured interviews with all patients in their home. Each debrief lasted between 53 minutes and 1 hour and 45 minutes. Before each interview we prepared screenshots of patients' dashboard views showing events of interest (e.g. worsening or improvement in measures between two days) and summaries of log data for each patient used to cue recall in the debrief. While patients filled out demographic questionnaires, one researcher scanned the diaries for events or activities of interest and prepared interview prompts.

The interview focused on following topics: COPD-related activities for managing disease, context of use and comparison with previous telehealth use. We showed a short video of THC to those patients, who had used it previously, to remind them (since they had not used it for a while) before we engaged in a discussion on comparing their previous telehealth experience with the system they had used.

Inspired by the grounded theory approach [23], we transcribed and coded the audio-recordings using an initial code list. We then defined emerging themes through an iterative

Patient ID	P1	P2	P3	P4	P5
Age	80	76	69	67	76
Gender	female	male	female	male	female
COPD severity	-	Moderate	Severe	Very severe	-
Years since COPD diagnosis	7	6	10	25	10-12
Living alone	Yes	No	No	No	Yes
Highest level for education	Higher education	Higher education	Higher education	Apprenticeship	Primary school
Technology experience	Daily	None	1-6 days a week	Daily	Daily
Supplemental oxygen	No	No	Yes	No	Yes
Vision problems	-	-	-	Color blind	Cataract
Other health-related conditions	Fibromyalgia, osteoporosis, diabetes	Liver problems	-	Epilepsy, heart problems	Connective tissue, diabetes
Telehealth experience	AF (3 months)	THC (6 months)	-	AF (3 months)	THC (6 months)

Table 1. Patient characteristics

process of reviewing the codes. We also include findings from the logging and diaries in the following.

Results

From the analysis, we identified five themes: (1) Barriers for reflection and system use, (2) using measures as health status indicators, (3) feeling empowered in everyday life, (4) questioning and gaining self-knowledge and (5) becoming motivated to self-improve. While some of the themes overlap, there are also noticeable differences related to different levels of reflection.

Barriers for Reflection and System Use

All patients mentioned the agreement with us as the main motivation for taking the system into use, except P3 who mentioned doing it, because she wanted to know her current status (*self-knowledge*) and what she could do about it (*self-improvement*).

Similar to patient types found in Study 2, passive patients only took the role as data providers when using the system and did not consider it their job to engage in reflective activity of the self-tracked data. We classified three patients (P1, P3 and P4) as active patients and two patients (P2 and P5) as passive patients.

Passive patients were not motivated to reflect on their own health, *"we can not do anything except measure"* (P2) and *"if the bright minds can not make sure that I get better, then neither can I do anything about it"* (P5). According to P5, reflecting on self-tracked data involved speculating about things that she did not believe she could change, *"I do not worry about things that I can not change"*, but she already engaged in a similar reflective activity in order to change things, e.g. using the pulse oximeter to e.g. adjust her supplemental oxygen level, when she did not feel well.

Frequency and average duration of system use were not indicative of whether patients were classified as active or passive (See Figure 5). All patients had used the system approximately 4-5 times, except one active patient (P4) who had used it nine times. Two of the active patients (P1, P4) logged in once during the trial period for the purpose of interacting with visualisations supporting long-term reflection alone.

All patients except P1 read and acted on reflective questions (e.g. system asked P4, *"You have several measures that are yellow or red. Have you explored what your measures might have been affected by? (Look into previous measures)"*, wherafter P2 followed the instructions and explored multiple measures and context data).

Patients used the system on average between 4 minutes and 51 seconds and 12 minutes and 30 seconds per session (M: 9.28), where session refers to logging in, interacting with the system and logging out. The longest session lasted 32 minutes (P4), consisting of entering measures, answering reflective question and interacting with visualisations. The shortest session lasted 3 minutes, consisting of only entering measures (P2 and P4). Both passive patients and one active patient (P3) used approximately 3/4 of the session time on the collection page.

Using Measures as Health Status Indicators

Patients attached importance to their subjective feeling (i.e. how they are feeling) and their engagement in reflective activities on symptom changes depended on whether they felt good or bad. When patients felt good, they did not necessarily see a reason to engage in the reflective activity about symptom change. Revisiting past using the provided time series graph was not of interest to the majority of the patients, which was also reflected in the low usage among patients (except P4).

Reflecting on the past was negatively charged by P1, *"that's not something I walk around and think about .. Life gets too strenuous if you walk around and think about that [bad days in past]"* (P1). Attaching importance to embrace good days were important to patients, e.g. *"[if] I actually feel good, then I do not worry about how I felt yesterday"* (P3). On the other hand not feeling well in the moment triggered reflection, *"if I do not feel like everything is fine, I might start thinking why (...) it depends on how I am feeling"* (P4).
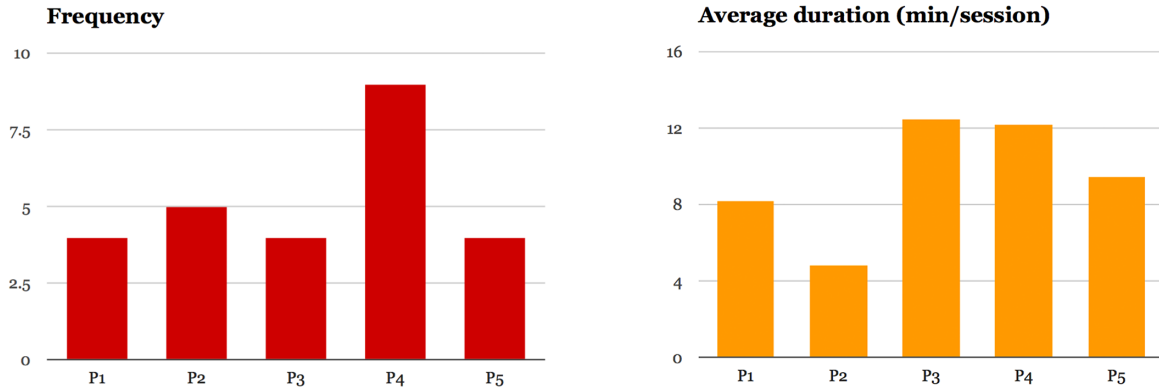
Figure 5. Frequency of use (left) and average duration of use in min/session (right)

Patients already used their pulse oximeter to check their current status, clarify whether oxygen saturation was the cause of not feeling well and then initiated action to improve their condition (e.g. breathing exercises if oxygen saturation measure was too low). The dashboard provided a similar indication of health status using multiple measures, *"it [dashboard] is a measure of one's symptoms (...) altogether it of course becomes how you are feeling"* (P1).

The self-tracking activity informed and increased the awareness of how patients were feeling, *"I start noticing three times a week, how am I feeling right now?"* (P4). P3 found it helpful to have it visualised on the dashboard what caused her not feeling well, *"you can not even go to the doctor and learn about your status and why (...) that you can do here [dashboard]"* (P3).

Questioning and Gaining Self-Knowledge
In using the system, some patients had gained insights and self-knowledge that they had not previously been aware of. Patients asked themselves questions, increasing their awareness on causes of symptom changes.

Reflective questions in combination with annotating measure with context variables in the system triggered reflection in P3, who had identified that weather had an impact on her breathing difficulties, *"(..) with dyspnea, I had not thought there could be other [reasons] .. I just had breathlessness, done. (..) suddenly I realized how much I was affected by the heat (...) it happened when I sat with the system and those questions asking 'why?'"*. Annotating with context variables supported evaluating different causal explanation *"I have started thinking about it (...) I think, 'no it's not that [stress]', 'Talk? No I haven't talked today' .. and then I think 'it's the weather'"* (P3).

P4 had started reflecting more on the day before and compared it with the presence to assess whether he could improve anything from yesterday, *"I become very conscious about, how did I feel yesterday? Do I also feel like that today?"* (P4).

Feeling Empowered in Everyday Life
Using the system had a transformative effect on P3, who had obtained a new perspective on her disease after using the system and an understanding of, what the measures meant. She felt that having COPD in many years had made her passive and lost hope on being able to do something.

"My memory has been stuck, so I thought that is just how it is.. You give up a little and get tired of it [COPD] (...) without doing anything about it, because nobody says anything.. but this [the system] does. It makes you aware about the situation (...) My doctor always told me that it is all because of my condition. The system makes me think that he is not right. I might have to make demands, then I might get better." (P3)

Some patients (2/5) felt more empowered in planning and overcoming everyday tasks. P3 mentioned not wanting to embarrass herself publicly and felt that she could plan to avoid such situations because of the gained knowledge about how she was affected by the heat, *"now I can make up my mind beforehand [whether to go outside in the heat], because I know how it will end, now when I have been told.."* (P3). Similarly, P4 found that it provided him with a feeling of safety knowing that he was within the recommended levels in terms of his measures, which he could see on the dashboard, *"I'm on the right track then. I do not have to worry about going to folk school or something else"* (P3).

Social Responsibility: Active patients who did not live alone (P3, P4) felt socially responsible towards their relatives. *"I can become unsure about how I am feeling.. (...) I do not want to expose my husband and daughter unnecessary [frightening events] it is about balancing..I learn more about that now, so that I do not expose them [relatives]"* (P3).

The self-tracking activity made P5 feel egocentric, but he considered it important in order to be able to do what was best for himself and his relatives, *"I have to be self-centred (...) I have to do things right for myself and in time, so that I also treat others right"* (P4).

Becoming Motivated to Self-Improve

Active patients (2/3) started setting goals themselves to improve measures. This involved seeking new knowledge, *"I've tried to acquaint myself with BMI because I wanted to have a goal to follow.. [because] I wondered about the arrows [in the system]"* (P4).

P3 had not previously been aware of the severity of her weight problems had become aware of the need of improving her condition, *"I have not thought about it before, but when you suddenly get it in writing (...) being confronted with it, I have to do something about it (...) it's for my own good"* (P3). She had started making changes to her eating habits in order to lose weight (eating less, thinking about what she eats, etc.) and mentioned being more aware of engaging in proactive behaviour, *"[less] coughing, that's about getting better at using the PEP device.. Not just saying, oh, you are running into a pneumonia, now you have to use it', it's about using it [PEP device] several times a day"*.

Some patients used the color indicators and arrows in combination as indicators of status and used them as goals, *"I want all of them to be green and that things are making progress"* (P3) and *(...) when the arrows are pointing down I assume it is not so good, that's the wrong way"* (P4). Seeing the progress in the system and that it paid off to change her behaviours, motivated P3 to ease off medication intake, which she had tried several times in the past without success. *"They had difficulties easing me off because I have had high doses for so many years (..) but this time I thought (...) now you have to stop (...) I did .. I needed some days and then it was over"* (P3).

Actionable Advice: Both P1 and P3 requested actionable advice on what they could do to improve their conditions. P3 needed advice on how to progress towards her goal of losing weight taking into account her other health-related conditions *"to get help when you also have diabetes, that would be nice"*. P1 mentioned needing actionable advice on improving measures, *"when you sit alone, have breathing difficulties, you cough and you have phlegm, you think, what can I do? It is the alpha and omega"*.

DISCUSSION

We investigated how concrete design decisions regarding entry and interaction with data affected reflection in telehealth among COPD patients. Our findings indicate that active patients benefited by becoming more informed and aware about their health status using the features in the system, leading to increased empowerment in their everyday life and motivation to self-improve by setting goals. In contrast, passive patients were not interested in reflecting on their self-tracked data, but the lacking motivation to reflect on data was not a barrier for their engagement with the system (e.g. only little difference in frequency and average duration of use between active and passive patients).

Previous literature shows conflicting results in terms of effectiveness of telehealth interventions for COPD patients on utilization of healthcare services and health-related quality of life [21]. We found that patients were highly motivated to log and submit data, because it provided them with sense of security to be monitored by a healthcare provider. However, we also observed an implication in current telehealth design, where some patients lacked knowledge or awareness on submitting reliable and valid measures, potentially hindering healthcare professionals in the early detection and initiation of treatment (the purpose of telehealth).

Our findings suggest that some patients did not want to be confronted with a past that cannot be changed, whereas visualising discrepancies that can be changed, encouraged some active patients to set goals in order to self-improve. Whether patients in fact become more self-managing requires a long-term study on behaviour change. While we provided patients with visualizations of past (history data) with the purpose of supporting reflection on change over time [24, 7] and patients mentioned benefits in seeing improvements in their measures, we found low usage of such visualizations in the real context, unless patients were curiosity-driven. Previous studies found that self-tracked data reminded people with a chronic health condition of negative aspects of their condition [14, 1], which could explain our findings in relation to passive patients. Personal Informatics literature often mention simple visualisation of history as the method for supporting reflection [15, 16, 24], but considering the negative effects it can have on people with a deteriorating chronic condition (already suffering from anxiety and depression), we propose that future research investigates design to support seeing the positive and its effects.

We found that designing for discrepancies (color indicators and arrows) and questioning (reflective questions in the system) triggered awareness and encouraged exploring different causal explanations (R2, higher level of reflection [10]) among some active patients. Whether this also makes it possible for patients to detect onsets of exacerbations earlier is yet to be investigated. It is further unclear whether outcomes will be the same, if a healthcare provider monitors the data as in real telehealth.

Supporting previous literature on conditions for reflection [2, 25], our findings show that passive patients were not open-minded or motivated toward engaging in reflection on their data, making it a barrier for them to benefit from the designed features. Additionally, both passive patients in the study had limited experience in interacting and using a tablet, which might have been another barrier. Fleck & Fitzpatrick suggest that people need a reason or encouragement in order to engage in reflective activity [10]. In order to motivate passive patients to take a more active role in their own health, it might be necessary to consider other interventional methods or provide further support in the system to intrinsically motivate passive patients.

Some patients were highly motivated to engage in reflection and even started setting goals to self-improve, using features in the system as goals. However, similar to Epstein et al., our findings suggest that lacking knowledge or support (e.g on how to improve measures) can be a barrier for self-improvement (action) [9]. While only three of the five participants were active and engaged in the reflective activity, the study might have been subject to self-selection bias. Partic-

ularly, the patient who gained most out of the self-tracking activity, had no previous experience with telehealth and thus no experience with the benefits of self-tracking (e.g. gaining self-knowledge). Patients might have been more reflective, because we primed them to reflect by telling them the purpose of the study and asked questions during the debriefs that prompted reflection. As mentioned by Isaac et al., reflection can be triggered when trying to externalize thoughts or feelings, e.g. in diaries. In our study, the use of diaries as a data collection method might also have fostered additional reflection [12].

CONCLUSION

Our study explores self-tracking needs in telehealth and how concrete design choices affects reflection among COPD patients. We investigated user needs and concerns critical for the effectiveness of telehealth interventions using a synthesis of literature on Personal Informatics and analysis of interviews with COPD patients using a state of the art telehealth solution. While patients generally felt taken care of, our findings show that some patients were not providing reliable and valid data, either due to lack of knowledge or awareness. Neither did the system support reflection on the self-tracked data or any follow-up action.

Interviews, workbooks and feedback sessions informed the redesign of the telehealth system. We designed and developed a prototype to support self-reflection among the patients and evaluated it in a two week field trial. Our findings show that using simple color indicators and arrows for visualising discrepancies and asking reflective questions in the system can be a first step towards increasing active patients' self-awareness on their health status. Future researchers should investigate, how to support and motivate patients (both active and passive) e.g. through actionable advice to meet their self-tracking needs.

ACKNOWLEDGEMENTS

REFERENCES

1. Ancker, J. S., Witteman, H. O., Hafeez, B., Provencher, T., Van de Graaf, M., and Wei, E. "you get reminded you're a sick person": Personal data tracking and patients with multiple chronic conditions. *Journal of medical Internet research 17*, 8 (2015).

2. Atkins, S., and Murphy, K. Reflection: a review of the literature. *Journal of advanced nursing 18*, 8 (1993), 1188–1192.

3. Brandt, C. L. Study of older adults' use of self-regulation for copd self-management informs an evidence-based patient teaching plan. *Rehabilitation Nursing 38*, 1 (2013), 11–23.

4. Choe, E. K., Lee, N. B., Lee, B., Pratt, W., and Kientz, J. A. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM (2014), 1143–1152.

5. Chung, C.-F., Cook, J., Bales, E., Zia, J., and Munson, S. A. More than telemonitoring: Health provider use and nonuse of life-log data in irritable bowel syndrome and weight management. *Journal of medical Internet research 17*, 8 (2015), e203.

6. Chung, C.-F., Dew, K., Cole, A., Zia, J., Fogarty, J., Kientz, J. A., and Munson, S. A. Boundary negotiating artifacts in personal informatics: Patient-provider collaboration with patient-generated data. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, ACM (New York, NY, USA, 2016), 770–786.

7. Cuttone, A., Petersen, M. K., and Larsen, J. E. Four data visualization heuristics to facilitate reflection in personal informatics. In *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice*. Springer, 2014, 541–552.

8. Das, A., Bøthun, S., Reitan, J., and Dahl, Y. *Design, User Experience, and Usability: Interactive Experience Design: 4th International Conference, DUXU 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part III*. Springer International Publishing, Cham, 2015, ch. The Use of Generative Techniques in Co-design of mHealth Technology and Healthcare Services for COPD Patients, 587–595.

9. Epstein, D. A., Ping, A., Fogarty, J., and Munson, S. A. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM (2015), 731–742.

10. Fleck, R., and Fitzpatrick, G. Reflecting on reflection: framing a design landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*, ACM (2010), 216–223.

11. Initiative, E.-P. *The Exacerbations of Chronic Pulmonary Obstructive Disease Tool (EXACT): Patient-Reported Outcome (PRO)*. Evidera.

12. Isaacs, E., Konrad, A., Walendowski, A., Lennig, T., Hollis, V., and Whittaker, S. Echoes from the past: how technology mediated reflection improves well-being. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 1071–1080.

13. Koff, P., Jones, R. H., Cashman, J. M., Voelkel, N. F., and Vandivier, R. Proactive integrated care improves quality of life in patients with copd. *European Respiratory Journal 33*, 5 (2009), 1031–1038.

14. Li, I., Dey, A., and Forlizzi, J. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2010), 557–566.

15. Li, I., Dey, A. K., and Forlizzi, J. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing*, ACM (2011), 405–414.

16. MacLeod, H., Tang, A., and Carpendale, S. Personal informatics in chronic illness management. In *Proceedings of Graphics Interface 2013*, Canadian Information Processing Society (2013), 149–156.

17. Mols, I., van den Hoven, E., and Eggen, B. Technologies for everyday life reflection: Illustrating a design space. In *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, ACM (2016), 53–61.

18. Muller, L., Divitini, M., Mora, S., Rivera-Pelayo, V., and Stork, W. Context becomes content: Sensor data for computer-supported reflective learning. *Learning Technologies, IEEE Transactions on 8*, 1 (2015), 111–123.

19. Oh, J., and Lee, U. Exploring ux issues in quantified self technologies. In *Mobile Computing and Ubiquitous Networking (ICMU), 2015 Eighth International Conference on*, IEEE (2015), 53–59.

20. Patel, R. A., Klasnja, P., Hartzler, A., Unruh, K. T., and Pratt, W. Probing the benefits of real-time tracking during cancer care. In *AMIA Annual Symposium Proceedings*, vol. 2012, American Medical Informatics Association (2012), 1340.

21. Pedone, C., and Lelli, D. Systematic review of telemonitoring in copd: an update. *Pneumologia i Alergologia Polska 83*, 6 (2015), 476–484.

22. Pianosi, P. T., Huebner, M., Zhang, Z., and McGrath, P. J. Dalhousie dyspnea and perceived exertion scales: Psychophysical properties in children and adolescents. *Respiratory Physiology & Neurobiology 199* (2014), 34 – 40.

23. Preece, J., Rogers, Y., and Sharp, H. *Beyond Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons, Inc., New York, NY, USA, 2001.

24. Rivera Pelayo, V. *Design and Application of Quantified Self Approaches for Reflective Learning in the Workplace*. KIT Scientific Publishing, 2015.

25. Rogers, R. R. Reflection in higher education: A concept analysis. *Innovative higher education 26*, 1 (2001), 37–57.

26. Rooksby, J., Rost, M., Morrison, A., and Chalmers, M. C. Personal tracking as lived informatics. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM (2014), 1163–1172.

27. Sorenson, E., and Brath, R. Financial visualization case study: Correlating financial timeseries and discrete events to support investment decisions. In *2013 17th International Conference on Information Visualisation*, IEEE (2013), 232–238.

28. Ure, J., Pinnock, H., Hanley, J., Kidd, G., Smith, E. M., Tarling, A., Pagliari, C., Sheikh, A., MacNee, W., and McKinstry, B. Piloting tele-monitoring in copd: a mixed methods exploration of issues in design and implementation. *Primary Care Respiratory Journal 21*, 1 (2011), 57–64.

29. Verdezoto, N., and Grönvall, E. On preventive blood pressure self-monitoring at home. *Cognition, Technology & Work* (2015), 1–19.

30. Whooley, M., Ploderer, B., and Gray, K. On the integration of self-tracking data amongst quantified self members. In *Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014-Sand, Sea and Sky-Holiday HCI*, BCS (2014), 151–160.