



Citation for published version:

Jones, S, Hue, W, Kelly, R, Barnett, R, Henderson, V & Sengupta, R 2021, 'Determinants of Longitudinal Adherence in Smartphone-Based Self-Tracking for Chronic Health Conditions: Evidence from Axial Spondyloarthritis', *PACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 1, 16. <https://doi.org/10.1145/3448093>

DOI:

[10.1145/3448093](https://doi.org/10.1145/3448093)

Publication date:

2021

Document Version

Peer reviewed version

[Link to publication](#)

© ACM, 2021. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies March 2021 Article No.: 16 <https://doi.org/10.1145/3448093>

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Determinants of Longitudinal Adherence in Smartphone-Based Self-Tracking for Chronic Health Conditions: Evidence from Axial Spondyloarthritis

SIMON L. JONES, University of Bath, UK

WILLIAM HUE, University of Bath, UK

RYAN M. KELLY, University of Melbourne, Australia

ROSEMARIE BARNETT, University of Bath, UK

VIOLET HENDERSON, University of Bath, UK

RAJ SENGUPTA, Royal National Hospital for Rheumatic Diseases, UK

The use of interactive mobile and wearable technologies for understanding and managing health conditions is a growing area of interest for patients, health professionals and researchers. Self-tracking technologies such as smartphone apps and wearable devices for measuring symptoms and behaviours generate a wealth of patient-centric data with the potential to support clinical decision making. However, the utility of self-tracking technologies for providing insight into patients' conditions is impacted by poor adherence with data logging. This paper explores factors associated with adherence in smartphone-based tracking, drawing on two studies of patients living with axial spondyloarthritis (axSpA), a chronic rheumatological condition. In Study 1, 184 axSpA patients used the uMotif health tracking smartphone app for a period of up to 593 days. In Study 2, 108 axSpA patients completed a survey about their experience of using self-tracking technologies. We identify six significant correlates of self-tracking adherence, providing insight into the determinants of tracking behaviour. Specifically, our data provides evidence that adherence correlates with the age of the user, the types of tracking devices that are being used (smartphone OS and physical activity tracker), preferences for types of data to record, the timing of interactions with a self-tracking app, and the reported symptom severity of the user. We discuss how these factors may have implications for those designing, deploying or using mobile and wearable tracking technologies to support monitoring and management of chronic diseases.

CCS Concepts: • **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**.

Additional Key Words and Phrases: personal informatics, self-tracking, healthcare, chronic conditions, rheumatic disease, smartphones, wearable devices, self-report, engagement, adherence

ACM Reference Format:

Simon L. Jones, William Hue, Ryan M. Kelly, Rosemarie Barnett, Violet Henderson, and Raj Sengupta. 2021. Determinants of Longitudinal Adherence in Smartphone-Based Self-Tracking for Chronic Health Conditions: Evidence from Axial Spondyloarthritis. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 1, Article 16 (March 2021), 24 pages. <https://doi.org/10.1145/3448093>

1 INTRODUCTION

Self-tracking refers to the practice of systematically recording information about various aspects of one's life. Driven by technological advancements such as the miniaturization of sensors and the development of apps and devices for recording data, the practice of self-tracking has grown from niche habit of data-enthusiasts [9] and

Authors' addresses: Simon L. Jones, University of Bath, Bath, UK, s.l.jones@bath.ac.uk; William Hue, University of Bath, Bath, UK, wz331@bath.ac.uk; Ryan M. Kelly, University of Melbourne, Parkville, Melbourne, Australia, ryan.kelly@unimelb.edu.au; Rosemarie Barnett, University of Bath, Bath, UK, rlb60@bath.ac.uk; Violet Henderson, University of Bath, Bath, UK, violet@bath.edu; Raj Sengupta, Royal National Hospital for Rheumatic Diseases, Bath, UK, rajsengupta@nhs.net.

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, <https://doi.org/10.1145/3448093>.

‘quantified selfers’ [16] to a mainstream activity that is becoming increasingly pervasive in healthcare settings. Self-tracking provides opportunities for easier, more frequent and more accurate collection of data that can support clinical decision-making and self-management of health conditions. Recent studies have highlighted the utility of self-tracking data in the management of chronic conditions such as Irritable Bowel Syndrome [17, 38], Parkinson’s Disease [51], HIV [76], Migraine [66] and ME/CFS [22].

Many tracking technologies enable automation of data collection with minimal user input. Sensors are capable of continuously monitoring factors such as physical activity and sedentary lifestyle [65], sleep quality [53], respiration [60], blood pressure [85], and glucose levels [24], providing that users adhere to wearing devices as required. While these forms of data can be collected automatically, patients must often *manually* self-report on symptoms, behaviours and experiences relating to aspects of their well-being [80]. A broad range of factors may impact patients’ inclination to manually enter data, but poor adherence to self-tracking can erode patients and clinicians’ ability to make decisions based on the data [87] and hinder effective management of a disease. Hence, understanding the factors that influence adherence is paramount for designing and deploying effective self-tracking technologies.

In this paper we explore the determinants of self-tracking adherence (defined as the extent to which the user’s action matches healthcare providers’ recommendation of recording data on a daily basis). We focus on adherence in manually self-reporting data via a tracking app, rather than adherence to using automated tracking tools (such as wearing physical activity trackers), which has received considerably more attention in prior literature (e.g. [23, 78]). We do this in the context of axial spondyloarthritis (axSpA), a chronic rheumatic disease that affects the spine, causing severe pain, stiffness and fatigue [71]. Like many other chronic illnesses, AxSpA requires ongoing self-management of symptoms to ensure quality of life. The impact of some symptoms can be measured using automated approaches such as physical activity monitoring, whereas others like pain and fatigue require manual data entry. AxSpA is also associated with symptoms that are problematic and not well understood (e.g. disease flares [71]), meaning that self-tracking can potentially help patients and clinicians to learn about these symptoms.

This paper reports two studies that provide insight into factors that determine adherence with manual self-tracking of data. Study 1 comprises a large-scale ($N = 184$), longitudinal (up to 593 days) deployment of the uMotif self-tracking app, which patients used to record symptoms and other data each day. From this study, we measured and correlated axSpA patients’ adherence to daily tracking against features of the data that they provide (e.g. the severity of their condition, time of data entry, data selection), as well as other available information about the patient and the devices that they use. Study 2 comprises a separate online survey of 108 axSpA patients with self-tracking experience, gathering data about the factors that subjectively influence their tracking adherence. The results from both studies allow us to identify and explain correlates and predictors of tracking adherence in axSpA, which can inform understanding of self-tracking behaviours more broadly.

We find that 27% of the variance in adherence between patients was explained by six factors: the age of the user; the types of devices used for tracking (smartphone operating system and concurrent use of a physical activity tracker); the severity of their condition (pain); the decision to track a problematic symptom (hot flushes); and the time of day at which users submit their data. Our qualitative findings overlap with results from statistical analysis of adherence factors, providing explanations for why adherence may be affected alongside additional reflections on the user experience of tracking devices. These findings point to specific user groups that may benefit from additional motivation and encouragement to engage in the continued use of self-tracking technologies. We discuss possible design implications, such as approaches for making better inferences about gaps in tracking data.

2 RELATED WORK

2.1 Adherence in Self-Tracking

The growing popularity of tracking technologies and apps has made it increasingly common for patients with different health conditions to engage in self-tracking. This involves the collection of personally relevant data that can inform self-management of a disease and clinical decision-making involved in diagnosis, monitoring and treatment [52].

Self-tracking offers the potential to address information gaps between infrequent clinical check-ups, where data gathering traditionally takes place [35]. Frequent data collection can enable a more complete picture of patients' symptoms, behaviours and disease activity. However, prior work has shown that users do not always adhere to consistent tracking [87]. Data generated by patients is often shown to be fragmentary, due to sporadic tracking behaviours [78, 87]. Many self-tracking technology users lose interest in tracking over time; lapsing and forgetting to track as frequently [28] and eventually abandoning or migrating between technologies [18, 20].

Previous work has discussed the importance of considering *adherence* for the accurate interpretation of long-term tracking data [77, 78]. Adherence provides a measure of data completeness, reflecting how closely users have followed a data collection protocol (e.g. recording data X times per day, or every X days) over a particular period [78]. Thresholds for determining adherent behaviour in self-tracking have varied in previous studies [83], often depending on the objectives of the analysis and types of insights that are sought from the data. Furthermore, the application of different adherence thresholds often results in different rates of data retention [78].

Poor adherence impacts the validity of activity monitoring [82] and infrequent or irregular logging can result in data that does not reflect participants' actual behaviour [86], raising concerns for clinical decision-making [87]. Incomplete data can also prohibit meaningful comparisons of data between different patients or across different time periods, hinder longitudinal insights relating to disease progression, and hamper conclusions of significance. Unreliable data also reduces users' trust in their self-tracking tools [6].

Although the literature has noted the potential to improve user adherence through design [55], few studies have the explicit goal of identifying factors that affect adherence in self-tracking. Moreover, previous research has typically focused on adherence regarding the use of devices that automatically log physical activity data, including commercial Fitbits [59, 88] and other wrist-worn trackers [23, 47, 89]. Comparatively few studies have examined adherence with manual self-reporting of data using smartphone apps, i.e. where users are required to self-report data on a regular basis. Prior work has argued that wearable activity monitors foster greater tracking adherence and are generally favoured by users when compared to manual tracking applications [81]. The additional effort required during manual data entry, which is both frequent and necessary in the self-management of chronic health conditions, makes it important to understand factors that may influence adherence.

2.2 Determinants of Adherence

A number of studies suggest that specific factors may play a role in determining adherence with self-tracking. Doherty et al. [23] reported a large scale analysis of over 100,000 participants' adherence to physical activity monitoring, finding that factors including age, sex, time of day, and season predicted adherence. However, their study focused on whether participants wore an assigned tracking device, rather than adherence with recording data about subjective health status.

Other factors associated with tracking adherence have been reported as secondary findings from a range of studies. For example, tendencies to lapse in self-tracking have been linked with differing motivations for tracking [42]; the perceived benefits of tracking being outweighed by the effort required [28]; and significant changes in health goals, physical abilities and personal circumstances [18]. Issues with poor design aesthetics, comfort, usability and the effort of device maintenance are also thought to influence adherence [28, 29, 42]. Other work suggests that mismatches between self-trackers' expectations and the actual capabilities of the technology

can affect adherence [18], alongside incompatibilities with user's lifestyle or self-perceived image [42]. Issues relating to device abandonment have been linked to age [43, 69], with rates of abandonment for physical activity monitoring technologies significantly lower amongst older adults than in younger people [50].

Factors influencing adherence have also been discussed in the literature on ecological momentary assessment (EMA), which is a research method that involves surveying people about their "activities, emotions, or other elements of their daily life multiple times per day" [84, p. 3]. EMA is akin to manual self-tracking in that participants are typically asked to enter data in response to a prompt that appears on their smartphone [84]. Failure to respond to the prompt leads to gaps in a person's data record, hampering data analysis.

Like the literature on self-tracking, studies have shown that a range of factors influence adherence to EMA. For example, demographic factors including gender [72] and age [63] influence adherence, and situational factors including time of day [61] and the user's location [10, 64] impact the frequency of responses. EMA has recently been adopted to study the lived experience of chronic health conditions [e.g. 61, 72], and studies have shown that people with health problems typically exhibit lower adherence with EMA than healthy controls [31, 63]. Condition severity can also influence EMA adherence. In a study of bipolar patients, Gershon et al. [31] found that adherence was lowest among participants with higher degrees of symptom severity and a greater number of lifetime suicide attempts. These studies collectively suggest that adherence is impacted by a range of factors, but it is not clear whether these factors will similarly influence adherence to manual self-tracking. Our study builds on and extends this literature by producing a better understanding of factors that predict adherence with manual tracking of data for the self-management of axSpA, a chronic health condition.

2.3 Axial Spondyloarthritis: A Chronic Rheumatic Condition

Axial spondyloarthritis (axSpA) is a chronic, inflammatory arthritis primarily affecting the spine and sacroiliac joints. AxSpA is thought to affect approximately 1 in 200 adults globally [70]. The initial symptoms of AxSpA are chronic pain and stiffness in the lower back. As the disease progresses, spinal mobility often deteriorates and the pain and stiffness worsen. Severe cases can result in fusion of the spine. These symptoms are accompanied by fatigue, poor sleep quality, and changes in mood. AxSpA can result in peripheral manifestations including uveitis (inflammation of the eye), dactylitis (inflammation of the fingers or toes) and enthesitis (inflammation of the sites where tendons or ligaments insert into bone). These factors have a significant impact on patients' quality of life, affecting their ability to perform common housekeeping tasks and maintain social commitments. Staying in work also becomes difficult, with the average retirement age of people with axSpA lying at 36 years [75].

The standard approach to managing axSpA consists of tailored exercise regimes and pharmacological therapy. Both have been found to slow disease progression when used consistently, but there is currently no treatment that can reverse the course of the disease. Continual monitoring of symptoms and ongoing self-management are therefore key to ensuring quality of life. As with other chronic conditions, mobile and wearable technologies can enable the self-management of axSpA by allowing patients to track information about their symptoms, reflect on what they have recorded, and potentially discuss relevant information with their clinician [3]. Lapses in tracking may present obstacles to these discussions, reiterating the importance of understanding factors that determine adherence with long term self-care.

In the present work, we study patients' adherence with self-tracking of axSpA to learn more about factors that influence self-tracking of chronic health conditions. As noted above, axSpA is a lifelong condition that requires consistent, long-term maintenance. This means that axSpA self-tracking has the potential to provide a large volume of data about factors that influence behaviour over a long period of time. In addition, axSpA has a broad array of symptoms and contributing factors which must be tracked. These factors encompass a range of data types and require a variety of approaches to tracking. Forms of data that may be useful to track in axSpA include daily subjective ratings of pain, fatigue, and stress; temporal data such as occurrences of short

term flares; physical activity and sleep data from wearable devices; and factors such as whether a person has taken medication or performed their recommended exercises. These elements are consistent with self-tracking of other chronic conditions [3, 49, 54]. Our research investigates how these and other variegated data sources may predict adherence with daily tracking, providing a better understanding of factors that result in improved rates of self-tracking.

3 STUDY 1: METHOD

Our first study investigated the role of specific factors in predicting adherence with daily self-tracking of axSpA. The study was conducted in collaboration with the Royal National Hospital for Rheumatic Diseases (RNHRD) in Bath, United Kingdom, a large public hospital that specialises in the treatment of rheumatic diseases, including axSpA. The study was part of a larger research project to gather data on axSpA patients' symptoms, activities, medications and self-management behaviours for the purpose of improving understanding of the condition and its treatment.

Participants in the study were asked to manually track their symptoms using a smartphone-based tracking application called uMotif and a variety of optional wearable devices for daily data collection. The uMotif platform has previously been used in a number of studies that have investigated smartphone-based self-tracking and self-management [5, 27, 40, 62].

3.1 Ethical Approval

All procedures received ethical approval from the lead University's ethical review board and the Research Ethics Committees (RECs) within the UK Health Departments' Research Ethics Service (NHS)¹ - Reference: 13/SW/0096

3.2 Recruitment and Participants

All participants in the study were registered patients at the hospital. They were recruited to take part during face-to-face consultations with health professionals (for example at their regular clinical check-ups or treatment sessions), and via flyers and notices distributed at the hospital site. Participants enrolled in the study through an online registration process that directed them to download and install the uMotif app on their iPhone or Android smartphone. Identical versions of the app were available for iOS and Android.

A total of 184 patients provided informed consent to be in the study and record data through the uMotif app. All participants had received a clinical diagnosis of axSpA and were under the care of a specialist consultant at the collaborating hospital. 126 patients identified as women and 58 identified as men. Participants' ages ranged from 22–85 years (*Mean* = 52.33 yrs; *SD* = 13.59 yrs). Figure 1 shows three histograms for the self-reported ages of participants at a) the onset of axial spondyloarthritis symptoms, b) clinical diagnosis of the condition, and c) consent and commencement of participation in the present study. Participants had diverse axSpA disease severity, as measured using the BASMI scale [1]. The BASMI is a standardised measure that ranges from 0–10, where 10 represents higher limitation of movement due to axSpA. Patients in our sample had an average BASMI score of 3.00 (*SD* = 1.85, *Range* = 0.16–8.00).

3.3 uMotif Smartphone Application

The uMotif smartphone app (Figure 2) allows users to self-report data for a total of 10 data facets. In our study, eight of these facets were pre-determined, having been selected by the lead axSpA clinician for the project on the basis of their relevance for understanding and treating axSpA. The eight predetermined data facets were daily measures of: pain, fatigue, flare status, anti-inflammatory drug use, recommended exercise adherence, mood, stress and sleep quality. The remaining two data facets were chosen at each patient's discretion. Each

¹<https://www.hra.nhs.uk>

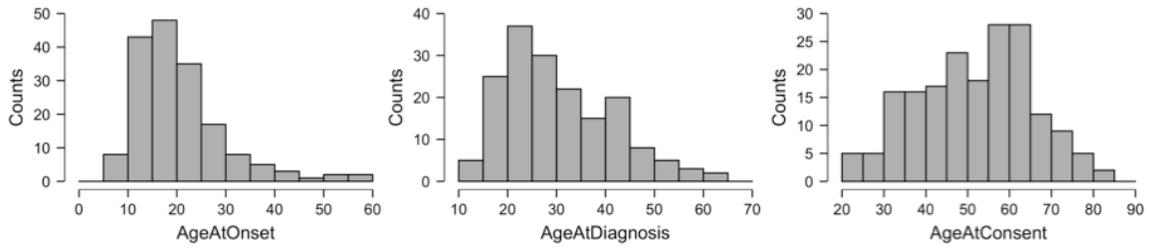


Fig. 1. Histograms for participant age at: a) symptom onset, b) clinical diagnosis, c) consent and participation in the study.

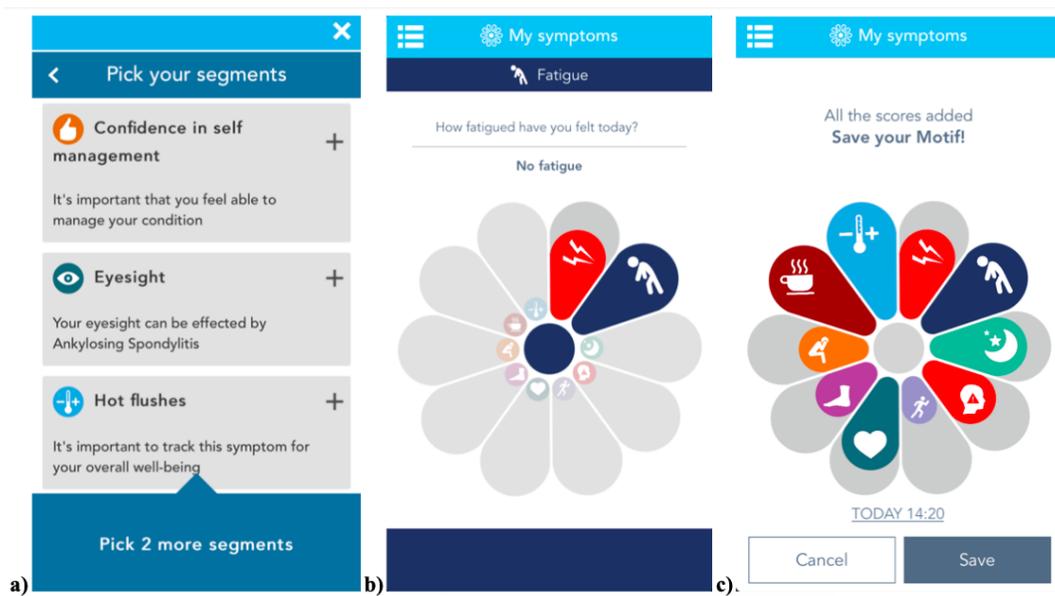


Fig. 2. uMotif app interface. a) Users select two data facets that they would like to track, from a list of 13 options, in addition to 8 data facets predetermined by AS clinicians. b) Data is recorded using the petal interface by dragging each segment to the desired value (e.g. rating fatigue from 'No Fatigue' to 'Very Fatigued'). c) A completed data entry for all data facets. Users can provide data as frequently as they wish, but are prompted to enter values at least once daily.

participant was able to select these from a list 13 options during app registration (see Figure 2a). Options for the two discretionary data facets were daily measures of: caffeine intake, hydration, drug adherence, confidence in self-management, chest pain, eyesight, screen time, hot flushes/sweats, menstrual cycle, painful eyes, smoking, flare of psoriasis, and blood in the stool.

The interface for recording data in the uMotif app appears as a flower-shaped visualisation (see Figure 2b) consisting of 10 'petals'. Each petal corresponds to a single data facet. Data values for each facet are recorded using either a 5-point scale (e.g. rating fatigue on a scale from '1. No Fatigue' to '5. Very Fatigued') or a binary Yes/No choice (e.g. indicating whether anti-inflammatory drugs have been taken). Users enter data by placing their finger at the centre of the flower and dragging outwards along each petal. The design utilises a visual

metaphor similar to that of previous self-tracking systems [19] where a full flower represents more healthy or optimum values (e.g. lack of symptoms and positive outcomes or behaviour) and a smaller flower conveys a less healthy or desirable status (i.e. worse symptoms). Users can only save data if all petals have a valid input, hence users were required to fill in data for all symptoms.

During the study, reminder notifications to complete the petal inputs were pushed to the user's smartphone at 8pm every day by default, if data had not already been entered. Users could opt out of reminders or alter the time and frequency settings of the reminders.

In addition, the uMotif app allows users to import data from wearable devices (e.g. Fitbit, AppleWatch, Garmin) as well as fitness applications and health data aggregators like MyFitnessPal, HealthKit and HumanAPI. These tools provide objective measures of exercise, sleep and diet, based on automatically collected data. Participants were invited to connect these services with the uMotif app at their own discretion. Amongst the 184 participants in our study, 82 participants connected physical activity tracking devices with the uMotif app. The remaining 102 participants did not connect any additional services to the uMotif app. All provided data was available to the researchers for analysis, in addition to the data from the petal interface.

3.4 Procedure

Each participant was asked to download and use the uMotif application as part of a rolling recruitment process that ran from the 5th April 2018 to the 21st November 2019.

Participants were told that although *daily* data entry was encouraged (i.e. entering self-report data once per day), any amount of information they could provide may be useful, including restarting data entry after inactive periods. The clinician managing the study initially asked patients to keep entering data for as long as possible.

Participants were informed that in addition to being able to view and reflect on their own data, their data may be reviewed by their clinician and used to support recommendations about management and treatment of the condition. They were also informed that aggregated data, from all of the participants in the study, may be used for research to understand the condition more broadly, for example by studying patterns and trends in symptom data.

Data entered into the app was stored in a secure database that was accessible to the clinician and members of the research team. The data analysed in this paper was downloaded on 21st November 2019 (593 days after the first participant downloaded and began using the app). All participants were able to continue using the application beyond the study end date.

3.5 Analysis

In addition to data entered by patients when using the uMotif app, our analysis included information about the types of device used by each patient (smartphone operating system and accompanying activity tracking devices), as well as demographic information that had been collected by the hospital at the point of enrolment. This included age at enrolment, age at disease onset, age at diagnosis, sex at birth, and the patient's BASMI score, indicating condition severity [1].

Our analysis includes data from all 184 patient participants. In our analysis, adherent tracking behaviour was defined as completing the uMotif petal data input at least once per day. This is in line with the instructions that participants received – to provide data using the uMotif app on a daily basis. Adherence is therefore calculated as a percentage, representing the number of days with logged data during the period from the participants' first data entry, up until the date at which the study was closed and data was downloaded for analysis. Hence, a participant with high adherence has tracked data on a frequent basis and sustained this tracking for a long period. Conversely, a participant with low adherence may have tracked minimally over the study duration, or only for a short period before abandoning the app.

We applied linear regression analysis in order to identify significant predictors of the adherence measure for each participant.

4 STUDY 1: RESULTS

On average, participants were enrolled in the study with the opportunity to self-track for 521.3 days, spanning the period from their first engagement with the app through to the time at which the data was downloaded and analysed by the research team ($SD = 119.1$ days, $Range = 106-593$ days). The mean duration (in days) between participants' first and last data entry (i.e. duration of active involvement in the study, per participant) was 240.9 days ($SD = 207.6$, $Range = 1-578$). The mean number of adherent days (days with logged tracking data) per participant was 135.0 ($SD = 163.4$, $Range = 1-563$). The mean adherence rate for tracking uMotif petal data was 26.4% ($SD=29.9\%$). The distribution for each of these measures is shown in Figure 3.

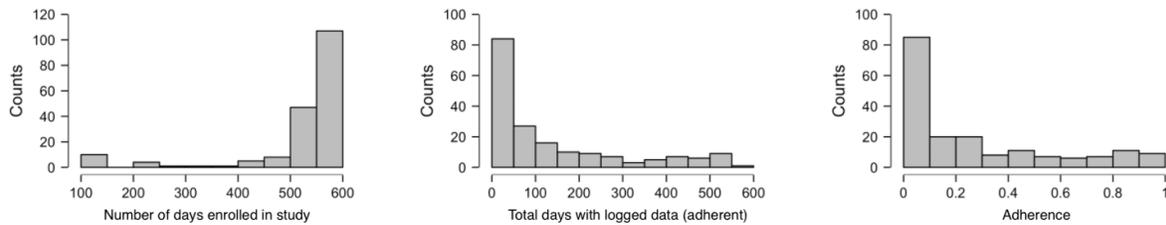


Fig. 3. Histograms for participants' duration of involvement in the study (from first data entry until the study closed), total number of days with logged data, and calculated *adherence*.

4.1 Linear Regression Models of Adherence

We applied multiple linear regression analysis to examine the relationships between self-tracking adherence and additional data variables outlined above. Table 1 illustrates variables in the regression models of adherence.

For our 'complete' model (see left side of Table 1), we used the 'enter' selection method, whereby all independent variables are entered into the linear regression equation at the same time. We present the results of this model for completeness and to allow for comparisons with a more parsimonious model, which explains the data with an optimal number of predictor variables.

We constructed an optimal regression equation (see right side of Table 1) using 'forward' variable selection, reducing the set of predictor variables to those that account for nearly as much of the variance as the complete set of all variables. Forward selection adds particular variables that give the single best improvement to the model one at a time, until no further predictors can significantly improve the model. Predictors were added in the following order: *Evening* ($\Delta R^2 = 0.115$; $\Delta F=22.112$, $p<.001$); *Age* ($\Delta R^2 = 0.072$; $\Delta F=14.946$, $p<.001$); *Hot flushes/sweats* ($\Delta R^2 = 0.034$; $\Delta F=7.312$, $p=0.008$); *Operating system* ($\Delta R^2 = 0.026$; $\Delta F=5.769$, $p=0.017$); *Uses physical activity tracker* ($\Delta R^2 = 0.03$; $\Delta F=6.879$, $p=0.01$); *Average pain* ($\Delta R^2 = 0.018$; $\Delta F=4.147$, $p=0.043$).

The optimal model shows that 27% ($R^2 = 0.295$, Adjusted $R^2 = 0.269$) of the variance in adherence can be explained by a set of six statistically significant predictors: age at enrolment ($\beta=0.256$; $P=<.001$); operating system ($\beta=0.195$; $P=.006$); use of a physical activity tracker ($\beta=0.197$; $P=.005$); frequency of evening entries ($\beta=0.307$; $P=<.001$); average pain level ($\beta=0.138$; $P=.002$); and selection of 'hot flushes/sweats' as a discretionary data facet ($\beta=0.158$; $P=.025$). Hence, uMotif app users were more likely to adhere to tracking when they were older, using an Android device, wearing a physical activity tracker to accompany the uMotif app, entering data in the evening, and choosing to track hot flushes. We discuss these findings in more detail in sections 6 and 7.

Covariate	Complete					Optimal				
	β (SE)	t	p	95% CI		β (SE)	t	p	95% CI	
				Lower	Upper				Lower	Upper
Intercept	-0.354	-1.671	0.097	-0.773	0.065	-0.507	-4.348	<.001	-0.738	-0.277
Demographics										
Age at enrolment	0.289	3.362	<.001***	0.003	0.01	0.256	3.797	<.001***	0.003	0.009
Age at disease onset	-0.131	-1.361	0.176	-0.01	0.002	—	—	—	—	—
Age at diagnosis	0.098	0.974	0.332	-0.003	0.008	—	—	—	—	—
Sex at birth - Male	0.057	0.712	0.477	-0.065	0.138	—	—	—	—	—
Devices Used										
Operating System - Android	0.193	2.582	0.011*	0.028	0.212	0.195	2.813	0.006**	0.036	0.204
Uses physical activity tracker	0.137	1.412	0.16	-0.033	0.197	0.197	2.876	0.005**	0.037	0.198
Uses Fitbit	0.052	0.591	0.555	-0.086	0.16	—	—	—	—	—
Time of Data Entry										
Morning	-0.148	-1.13	0.261	-0.594	0.162	—	—	—	—	—
Afternoon	-0.148	-1.173	0.243	-0.632	0.161	—	—	—	—	—
Evening	0.082	0.486	0.628	-0.254	0.42	0.307	4.641	<.001***	0.179	0.444
Condition Severity										
Average pain	0.148	1.568	0.119	-0.016	0.136	0.138	2.036	0.043*	0.002	0.11
Average fatigue	0.062	0.568	0.571	-0.05	0.091	—	—	—	—	—
Average sleep quality	0.063	0.748	0.456	-0.043	0.096	—	—	—	—	—
Average mood score	-0.031	-0.282	0.778	-0.086	0.065	—	—	—	—	—
Average stress	-0.059	-0.605	0.546	-0.1	0.053	—	—	—	—	—
BASMI	-0.077	-0.88	0.38	-0.041	0.016	—	—	—	—	—
Variable Selection										
Caffeine intake	0.063	0.686	0.494	-0.072	0.149	—	—	—	—	—
Hot flushes	0.162	1.987	0.049*	7.834e-4	0.299	0.158	2.38	0.018*	0.025	0.268
Exercise Adherence	-0.06	-0.715	0.476	-0.163	0.077	—	—	—	—	—
Screen time	0.028	0.356	0.722	-0.107	0.154	—	—	—	—	—
Confidence in self management	0.08	0.916	0.361	-0.065	0.177	—	—	—	—	—
Hydration	0.056	0.643	0.521	-0.074	0.146	—	—	—	—	—
Chest pain	-0.126	-1.583	0.116	-0.227	0.025	—	—	—	—	—
Flare of psoriasis	-0.076	-1.035	0.302	-0.357	0.112	—	—	—	—	—
Smoking today	-0.116	-1.559	0.121	-0.425	0.05	—	—	—	—	—
Blood in stool	0.028	0.388	0.699	-0.259	0.385	—	—	—	—	—

Table 1. Linear regression model for predicting daily self-tracking adherence with the uMotif smartphone app. Left side = Complete model, Right side = Optimal model.

5 STUDY 2: METHOD

Our second study used an online survey to acquire a qualitative understanding of existing usage of self-tracking tools among the wider axSpA patient community. We used the survey to provide further context and explanation for the correlates of self-tracking adherence that were revealed in our analysis of Study 1. Furthermore, we use the patients’ accounts of their lived self-tracking experiences from the survey to unveil additional determinants of adherence that were not uncovered by the uMotif study.

Our survey collected information across four topics. First, *patient demographics* (e.g. age, gender, occupation, disease status and history). Second, *attitudes towards self-tracking* (e.g. views on tracking technology, tracking motivation and goals - including questions such as: “What is your general attitude towards self-tracking and

self-tracking technologies?” and “*What motivated you to start self-tracking?*”). Third, *tracking behaviours* (e.g. types of tracking technologies used, data collected, tracking frequency, tracking facilitators and hindrances - including questions such as: “*What makes you more/less likely to keep tracking?*”, “*When are you most likely to use tracking tools to self-report data?*”, “*Why did you stop tracking with this tool?*”). Finally, *data usage* (e.g. “*How useful/valuable is self-tracking of the following data types?*”). We collected this information using a combination of multiple choice questions, likert ratings and open-ended questions.

5.1 Ethical Approval

The study received ethical approval by the lead University’s Research Ethics Approval Committee for Health - Reference: REACH EP 1819/078. NHS ethical approval was not required, since survey participants were not recruited through the National Health Service.

5.2 Participants and Recruitment

We recruited participants from online patient communities. The survey was hosted at *onlinesurveys.ac.uk* for a period of 6 months and was open to all patients who have been diagnosed with axSpA and have had some prior experience of using a self-tracking technology.

We received responses from 108 participants (43 men, 65 women). Participants’ ages range from 22 to 77 (*Mean* = 49.38, *SD* = 13.38). 54 reported having no other chronic conditions besides axSpA and 54 reported having one or more additional chronic conditions. 49 participants reported using tracking tools that required regularly entering some self-reported data into a smartphone app (e.g. manually logging symptoms in apps such as MySpA, Chronic Pain Tracker, My Pain Diary and uMotif), with the remaining participants having used automated tracking (e.g. through wearing physical activity monitoring devices such as Fitbit, Garmin, Apple Watch and Samsung Galaxy Watch).

5.3 Data Analysis

We analysed the survey data via thematic analysis [11], using a combination of deductive and inductive coding. The first three authors of this paper began by reading the data to acquire familiarity. The second author then deductively coded the survey responses to label data that aligned with significant predictors identified in Study 1, allowing us to acquire a deeper understanding of why these factors impact adherence. The coder identified relevant data around three initial themes: condition severity, variable selection, and timing / routines. There was no data relevant to age or device type. In the next stage, the same author inductively coded to search for additional factors. This led to a further five initial themes: usability, effort, maintenance, battery life, and emotional impact. The first and third authors then refined the eight themes, consolidating emotional impacts with condition severity and classifying the other inductively-derived themes as user experience issues. This process ended with a final set of four high-level themes. The first and third authors rated their agreement with the assignment of each coded data point to the final four themes, resulting in an inter-rater reliability of Kappa= 0.934 (95% confidence interval: from 0.861 to 1.00). This is considered as almost perfect agreement [41].

6 STUDY 2: RESULTS

Our thematic analysis identified factors that participants subjectively described as affecting their adherence and engagement in self-tracking. These were *condition severity and emotional impacts*, *tracking individually relevant data*, *establishing routines and being reminded*, and *user experience issues*. The first three themes overlapped with factors exposed by Study 1, and the fourth was an additional factor that participants mentioned as important but which was not reflected in our model.

We did not encounter any responses to support or explain our quantitative findings relating to relationships between age and adherence, or operating system and adherence. This is somewhat expected, since respondents had seldom engaged in self-tracking for such a duration that they might reflect on how their adherence changed with age. Similarly, most respondents were unlikely to have engaged in self-tracking with multiple devices (e.g. smartphones with different operating systems) beyond their primary device.

6.1 Theme 1: Condition Severity and Emotional Impacts

Study 1 suggested that when people feel worse—specifically when they are experiencing more severe pain—they are more likely to adhere to daily self-tracking. Our qualitative analysis provides further insight into this trend. We found evidence that people become more inclined to track when their condition deteriorates: “*I tend to use it when I’m struggling*” (P58), “[*I track*] when I am feeling overwhelmed by this disease” (P64). Likewise, we found examples of respondents reporting that their motivation to track *decreases* when they are feeling healthy: “*I have not really used [the MySpA tracking app] consistently because I have so many other things to do and my AS is under control...*” (P2).

In some cases, the unexpected nature of changes in their condition was a more important factor than whether the condition was improving or deteriorating. For example: “*If symptoms or changes are regular or expected, [there is] no need to keep tracking with the app*” (P12). This points toward the occasional use of self-tracking tools as mechanisms for recording and monitoring disease fluctuations that are especially problematic, or which cannot be easily understood or explained. Hence, for these people self-tracking involves episodes of engagement and disengagement, and their condition status was a causal factor of adherence.

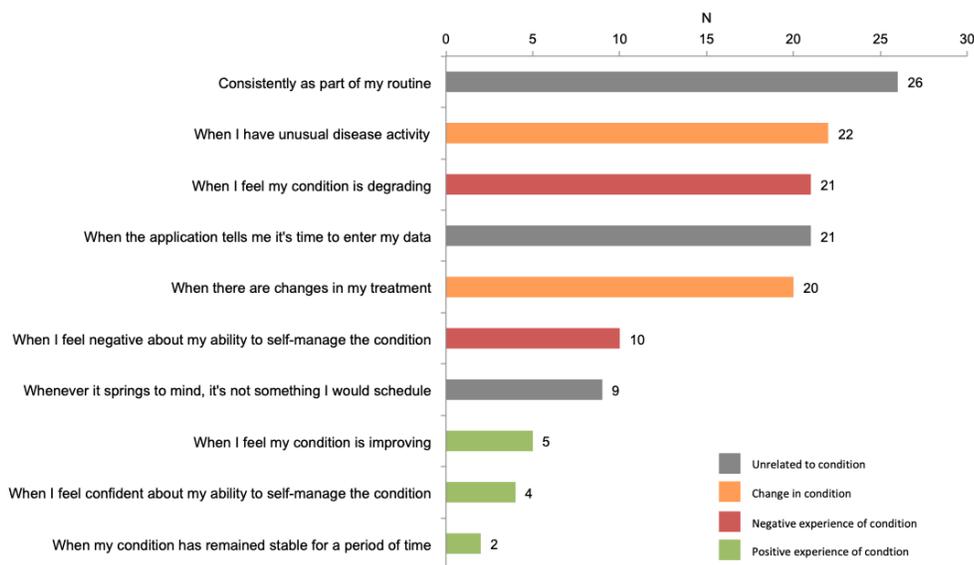


Fig. 4. “When are you most likely to use tracking tools to self-report data?” (respondents were able to select multiple answers)

To provide further insight into the subjective effect of condition severity on adherence, Figure 4 shows responses to the survey question “When are you most likely to use tracking tools to self-report data?”. This question was only asked of respondents with existing experience of tracking self-reported data (49/108). Respondents were able to

select multiple answers. While the majority of respondents (53%) reported using tracking tools consistently as part of their daily routine, 45% stated that they were more inclined to engage with self-tracking in response to unusual disease activity, and 42% reported that they had engaged in tracking when their condition was getting worse. Figure 4 also illustrates a general trend that positive experiences of a condition (shown in green) are less likely to drive self-reporting of data, as compared to negative experiences (red), fluctuations or changes in condition (orange), or factors unrelated to the condition (grey) such as tracking as part of a routine, or in response to prompts/reminders.

Although worsening disease symptoms were reported to drive adherence, 10% of respondents claimed that they were more likely to track when their condition was improving. P72, for example, preferred to record data when they were feeling well, as a way of “*continuing the belief my condition is under control*”, but was demotivated to track by “*apathy, anxiety and the feeling my condition is deteriorating*”. Mirroring these comments, other responses suggest that the relationship between condition severity and tracking adherence is not monotonic, i.e. it is not always true that as severity increases, so too does adherence (as suggested by our linear model). Some respondents reported that they stopped engaging with self-tracking when their symptoms had deteriorated to an unmanageable level, because it physically impacted on their ability to engage with the tool: “[*I’m less likely to track*] when I have a flare up” (P62); “[*I track*] when I’m feeling well enough to use tools. If I’m feeling too unwell I will not have the energy to use [them].” (P65); “*Sometimes I forget to log in my symptoms, especially if I’m fatigued*” (P106). These claims are similar to the findings of Matthews et al. [48], who suggested that patients with bipolar disorder were less likely to continue tracking at times when symptoms were most extreme.

We also found that emotional impacts of tracking, linked to condition severity, appeared to affect respondents’ adherence. Participants reported that they were worried about the negative feelings caused or amplified by the practice of self-tracking, and reflection over the self-tracking data itself. While reflection is often touted as a useful outcome of self-tracking [15, 21, 45], it risks provoking negative outcomes from focusing on the problems associated with one’s condition. P8, for example, said that they “*stopped tracking as it was all very negative. I filled it out when I got up in the morning and it constantly made me focus on pain so I stopped it. It made me focus on the negative aspects of my condition on a daily basis which was soul destroying*”. Another respondent said that “*it becomes depressing to analyse how I’m feeling every day—prefer to just cope as best I can and log when I think to*” (P58). P41 felt that looking back at data could have a negative impact “*I believe strongly that to cope with the demands of a chronic painful condition one needs to operate in their here and now, rather than carrying a sense of the ongoing pattern of limitations pain and tiredness.*” For P62, prompts to increase physical activity were incorporated into their self-tracking tool (Fitbit). This had a negative impact on their adherence when their condition affected their ability to be active: “[*I’m less likely to track*] when I have a flare up. I feel guilty when I get reminders to try get my steps goal, and I can’t because I’m in too much pain” (P62). These findings overlap with previously reported adverse effects of goal-driven self-tracking for certain chronic health conditions [22].

Developing positivity around the condition and supporting feelings of being pro-active were further reasons to keep tracking: “*It helps me feel that I am doing the right thing*” (P26). One participant said that it gave them “*Hope and encouragement*” (P31) and another gained “*Satisfaction at completing my daily targets and psychologically it feels like I am achieving something positive for my mental and physical health.*” (P49). However, others said that feeling despondent about the purpose and outcomes of self-tracking would affect their adherence: “*When you have done everything you need to do and AS still flares, it makes you think, what’s the point?*” (P78). Some participants discussed experiencing emotions if they did not reach certain activity thresholds: “*I feel guilty now if I don’t reach my daily step target.*” (P38).

6.2 Theme 2: Tracking Individually Relevant Data

The second factor in participants' responses describes how the type(s) of data being tracked impact adherence. Respondents described tracking more frequently when collecting types of data that they considered to be "relevant" (P14), "useful" (P50, P65, P86, P94) or having "value" (P76). Being asked to record data that did not have clear relevance to their condition was a barrier to adhering to regular tracking: "The [self-report] questions don't always seem relevant so it's difficult to stay motivated. I'm putting in a lot of information that does not appear relevant, but [there's] no tracking of painkillers in my app" (P45). P14 wanted to: "modify it [the tracking application] so I only have to answer questions that are relevant to me", suggesting that having individual control over the data types to be tracked may improve adherence.

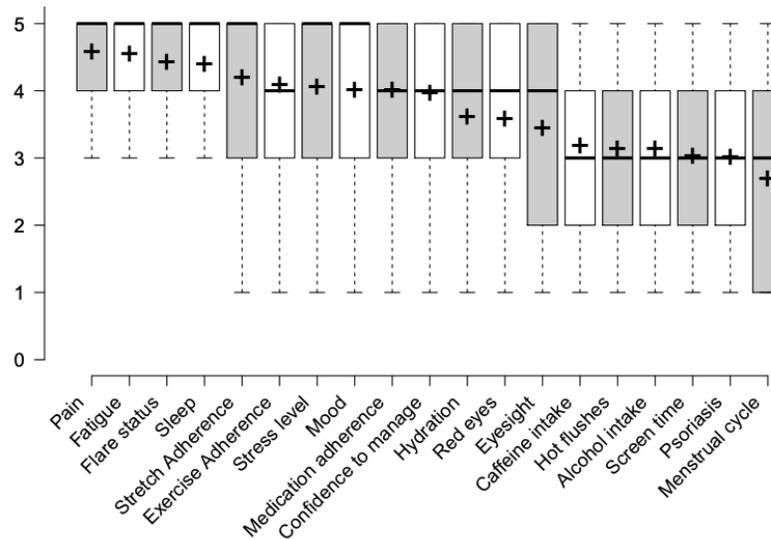


Fig. 5. "How useful/valuable is self-tracking of the following data types? (from 1 - Not at all useful to 5 - Extremely useful)" '+' indicates Mean. Boxes indicate Median and the Interquartile Range (IQR). Whiskers show lower and upper 1.5*IQR.

Our survey asked participants to reflect on the most useful or valuable data type that their devices allowed them to track (see Figure 5). The most common response was *pain*, followed closely by *fatigue*, *flare status* and *sleep*. For many, managing pain is their primary objective for tracking: "It has pushed me to ask questions about the variation in pain level in one of the joints affected" (P1), "I was not understanding my constant fatigue and pain and not knowing what helped or made it worse" (P99), "[I started tracking] to understand my pain and fatigue, was there a correlation?" (P85). These responses link with our statistical finding that higher adherence is associated with increased pain. Pain is the hallmark feature of axSpA and, as such, when it is problematic patients may feel inclined to track this symptom as a way of understanding it more deeply.

Our statistical model also revealed that the decision to track 'hot flushes/sweats' within uMotif was a significant correlate of higher adherence. Responses relating to this symptom's impact on adherence were not prominent in the survey data. However, one participant reported that for them this symptom provided some predictive power, alerting them to a potential oncoming flare in their condition: "Night sweats rather than hot flushes seem to signal a flare is coming" (P21). Hence, they viewed this as a particularly important data facet to track.

Responses also suggest that rather than all data types having uniform value to patients, particular data types may provide greater utility to a small proportion of patients. This may be when the symptom is particularly

problematic, or offers some prognostic capability. Our results suggest that increased adherence depends on data having relevance and value for the specific individual user, not just the condition in general. Respondents valued having a tool that was “*tailored to my need*” (P1), “*customisable*” (P14), and allowed “*multiple tracking options*” (P59), to “*track everything I need it to*” (P105).

6.3 Theme 3: Establishing a Routine and Being Reminded

Respondents suggested that establishing a regular tracking routine was an important component of adherence: “*You do need to discipline yourself to try to enter the data on a daily basis*” (P43), “*It is a conscious effort and needs to be incorporated into daily routines before it will become natural*” (P79). However, concerns were expressed about difficulties with *remembering* to track. Nine participants noted that their own forgetfulness influenced adherence, with P55 noting that they “*sometimes forget*” and P14 saying “*it’s very easy to do, just sometimes forget to do it*”. Similarly, P56 and 58 noted: “*Providing the data is relatively effortless but remembering to do it is more of an effort*” (P56), “*I don’t remember to log every day... I log when I think to*” (P58).

Our earlier statistical analysis revealed that participants who were entering data into the uMotif app during the evening hours were more likely to be adherent. This time coincided with the app’s default reminder notification and aligns with the survey responses, which suggest that reminders and notifications can drive adherence. Many participants reported that alerts, alarms and reminders help them to remain adherent: “*Diaries require routine and habit. Apps however have reminders... that is helpful*” (P64), “*Alarm beeps on phone, I fill in the questionnaire*” (P104).

One respondent reported that the presence of their wrist-worn tracking device helped them to remember: “*The watch itself is a reminder*” (P40). Our statistical analysis revealed that those who used physical activity tracking devices alongside the uMotif app for manually reporting data were more adherent, suggesting that attending to devices that enable automated tracking (e.g. capturing physical activity) may also stimulate engagement with manual forms of tracking.

6.4 Theme 4: User Experience Issues

Participants described four user experience issues that affected their tendency to self track. These issues were related to the design of devices and how these placed demands on the user when entering data.

The first was related to *app / device usability*. Nineteen participants explicitly mentioned that “*ease of use*” and “*simplicity*” were qualities that made them more likely to keep tracking. Likewise, poor user interfaces, complicated functionality and tools that were “*difficult to use*” were reasons to give up tracking. Some participants also mentioned device comfort as something that could affect their adherence.

A second issue concerned the *amount of effort* required for logging data. Activities that involved protracted data entry were generally regarded as costly, compared to automatic tracking and passive data collection: “*I only collect data from my watch. I wear the watch daily, no real effort on my part*” (P40). Manual data input was noted as “*time consuming*” (P91) and could affect adherence in combination with a poorly designed user interface: “*Some apps require a lot of self-entry which takes a lot of effort. Fitbit app is easy because it tracks it for you via the watch*” (P86). Users may also become bored or fatigued by the repetitive nature of data entry, suggesting a need for more engaging techniques for keeping users interested [55]. P86, for example, expressed frustration with “*having to enter in results all the time in order to see any useful data*”. Setup costs were also off-putting. P36 claimed that they were less likely to keep tracking “*if a lot of effort is required to set it up each day*”.

In addition to the effort of data-entry, the *maintenance* of devices may influence adherence. Five participants mentioned that “*poor battery life*” was a source of frustration and reason for device abandonment. Previous work has highlighted that users often forget to take devices with them or forget to charge batteries and that this can impact engagement [8].

Finally, the *motivational features* of tracking technologies may affect adherence by altering users' incentives to track. For example, some apps employ social comparison to benchmark users' performance and encourage competition. Our survey showed this was a double-edged sword. Patients who commented on their use of social features sometimes viewed these as positive: "I'm in competition with my 2 daughters so that encourages me to use it" (P81). Others noted that social comparison could lead to "Feeling like I've achieved less than other people" (P31) or "feeling guilty if I don't reach my daily step target." (P38), which demotivated them from tracking.

7 DISCUSSION AND IMPLICATIONS

The aim of this paper was to explore factors that predict adherence with manual self-tracking by patients with axSpA. In Study 1, we analysed 184 patients' use of the uMotif app for up to 593 days. We found that six factors were significant predictors of adherence, namely: entering data in the evening; the age of the user; selecting to track 'hot flushes/sweats'; smartphone operating system; use of a physical activity tracker; and mean average pain rating. These factors account for 27% of the variance in measures of adherence. The findings of Study 2 overlap with the results of our statistical analysis, providing further insight into the reasons why these factors may be influential and also noting that user experience issues with tracking devices may further impact adherence.

Overall, these findings provide a deeper understanding of the factors that impact manual self-tracking adherence for patients with a chronic health condition. Prior work has investigated factors associated with the use of physical activity tracking devices, (e.g. [23]), yet few studies have reported on adherence to manually providing self-reported data within a smartphone app. Our work develops further understanding of tracking adherence, which may support health professionals in identifying factors that can help or hinder the collection of useful data. Our analysis also enables assessment of the relative impact of different adherence determinants. Future work may wish to account for these factors, for example by developing additional interventions and guidance to maximise adherence and target those who are less likely to adhere. In the following sections we discuss each of the factors in turn, suggesting possible implications for those who wish to design, deploy or use self-tracking applications in similar contexts.

7.1 Age and Adherence

Our Study 1 analysis revealed that participants' age was a significant predictor of their adherence, accounting for 7.2% of the variance. Specifically, older users were more likely to adhere to manual self-tracking using the uMotif app. This result is in line with age-related differences in behaviours, attitudes and intentions associated with self-tracking, which have been reported in prior work. For example, older adults are more likely to judge activity tracking devices to have greater long-term utility [37, 50]. McMahon et al. [50] found that older users are less likely to abandon tracking technologies, such as physical activity monitors, compared to younger people. [63] reported that higher age was related to better compliance to EMA prompts. For example, people aged 30 years had a compliance of 76% compared with compliance of 86% for those aged 60.

Jaana and Pare [37] reported that the majority of apps used by older adults in Canada are health-related, and that older adults are more likely to sustain long-term use of these apps. These results are despite a broad emphasis on tracking technologies catering to the needs, interests, goals, and practices of younger users [13].

Health issues and disease management typically become more salient and important for individuals as they get older [25], which may lead to an increase in health-consciousness and decrease in reticence to undertake health actions. For instance, when symptoms are perceived as more serious, the response to seek treatment is quicker in older adults than in young adults [44]. The results of our Study 1 suggest that age-related shifts in health attitudes and behaviours may also lead to an increase in adherence to self-tracking.

Prior work has also demonstrated that patient-clinician relationships are subject to age-related differences. Peck [56] found that clinicians were more likely to facilitate "*patient-centered encounters*" – where "*patients*'

desires and expectations are incorporated into the medical decision-making process” [57] – with older adults over the age of 65. Age has also been shown to influence the style of interaction between patient and clinician, as well as overall satisfaction with the interaction. We hypothesise that this relationship could be an underlying factor in the increased self-tracking adherence observed in the present study. Patients were encouraged to self-track by their clinician, therefore those with greater “trust, knowledge, regard, and loyalty” [14] in this relationship may have been more receptive to this advice and adherent to instructions. Similar effects of trust in one’s clinician have been shown to be associated with better adherence to therapy using pharmaceutical drugs [79].

Finally, we posit that age-related differences in lifestyle may contribute to the observed variations in adherence. Older people spend their time differently than their younger counterparts, with notable reductions in leisure and social time [46]. Since tracking was often reported as a burden and time commitment in our survey responses, available ‘free’ time may have an impact on users’ engagement with a tracking app.

An implication of this finding is that younger users may need additional motivation and prompting to engage with self-tracking tools in order to foster adherence, despite evidence that they have higher levels of smartphone usage and are often perceived as being more experienced users of technology [33].

7.2 Data Selection and Adherence

Our results reveal that users who selected to track ‘hot flushes’ exhibited higher adherence than those who did not. Hot flushes, i.e. feelings of intense heat, feverishness and experiencing sweats, in axSpA are not well understood, but have been described by patients as highly unpleasant and amongst the worst symptoms².

While there is no clear explanation for the observed association between selecting to track hot flushes and adhering to self-tracking more generally, it is possible that the presence of this symptom corresponds with a worse experience of the disease (in line with our earlier findings relating to condition severity). Those who suffer hot flushes/sweats (and therefore choose to track them) may be more inclined to track because of their negative disease experience. Moreover, it is also possible that for some patients, this symptom offers prognostic capabilities, providing them advance warning of an oncoming flare in their condition. However, given that this is not a widespread trait, our results point chiefly to the importance of designing future tracking apps such that they enable users to customise which data types to track. While there is general consensus that tracking of the most predominant symptoms (e.g. pain, fatigue) is valuable, our findings suggest that users may also find unique value in additional data that speaks to their individual condition and experiences of a disease, which in turn elicits adherence.

7.3 Condition Severity and Adherence

In line with the previous indication that a worse disease experience leads to higher adherence, our model reveals that 4.3% of the variance in adherence scores can be explained by differences in the severity of participants’ pain. Users with higher mean average pain levels (based on daily likert scale ratings) were more likely to adhere to self-tracking. However, results from Study 2 also suggest that it is not consistently true that the worse a condition gets, the more likely a patient is to track. Some patients reported passing a threshold, beyond which they felt so debilitated by their disease that adherent tracking became more difficult to achieve.

These results are interesting when compared with prior findings regarding the impact of severity and lived experience of a condition on engagement with tracking. In Seppen et al. [67], rheumatoid arthritis patients reported that pain positively mediated their usage of self-monitoring app. That is, more pain promoted app usage and less pain discouraged app usage. A lack of symptoms on which to report was identified as a barrier for usage.

In contrast, Ancker et al. [2] explored patients’ perceptions of self-tracking and discovered that tracking is sometimes viewed as a form of ‘illness work’ which carries a significant burden. They found that self-tracking

²<https://forums.spondylitis.org/ubbthreads.php?ubb=showflat&Number=275597>

has potential to provoke strong emotional responses and that some patients experience negative feelings when seeing data values that remind them that they are unwell, or which undermine confidence in an ability to manage their disease effectively. Furthermore, some patients feel “*anger*” and “*injustice*” [2] when a technology serves as a reminder that health must be a significant focus of their attention on a regular basis. Our work uncovered similar emotional responses.

Overall, these findings suggest that people sometimes stop tracking because they are in good health, then begin tracking again when they were ill and have symptoms they want to record. They may later stop tracking again, either because their condition returned to normal or because it became so bad as to be unmanageable. Chiefly, these lapses create gaps in the tracking record, but it may be unclear as to why. It may be possible to develop intelligent predictions that ‘fill in’ gaps in the data on the basis of other available inferences, such as step count. For axSpA, reductions in physical activity recorded automatically by a wearable tracker may be a sign of immobility caused by a flare. The accuracy of filling in these gaps can be improved the more we understand about why people track, and how they feel when they are most likely to engage in tracking.

7.4 Smartphone Type and Adherence

Our model revealed that using the Android OS to log data with uMotif was associated with higher adherence, despite no discernible differences in the functionality of the app between Android and iOS. Previous studies have revealed differences in tracking behaviours between users of different smartphone platforms. Harari et al. [34] reported that Android users collected more sensor data and EMA responses per day than iOS users, but primarily attributed this to technical constraints on background processes running with iOS impacting the frequency of data collection. We note that such constraints did not apply within the present study. Pryss et al. [58] reported significant demographic differences between Android and iOS users of the ‘Track your tinnitus’ mhealth platform, with Android users being significantly older and having a significantly longer self-reported symptom duration than iOS users. We observed similar age differences in our study, with Android users (Mean age = 55.86, $SD = 12.8$) being 5yrs older, on average, than iOS users (Mean age = 50.65, $SD = 13.7$) ($t = -2.471$, $p=0.014$).

Shaw et al. [68] explored differences in the characteristics of Android and iOS users more broadly, noting similar demographic differences. They found that iOS users were more likely to be female and younger. Götz et al. [32] found significant differences in the wealth of users, with iOS users tending to have more financial resources, and in measures of the openness personality trait, with Android users tending to be slightly more open. These findings suggest that different groups of users may be attracted to particular types of devices. Further work may wish to explore the impacts of latent factors such as personality and attitudes towards technology on adherence in self-tracking, since these have rarely been accounted for in prior studies.

7.5 Physical Activity Device and Adherence

Our model from Study 1 also reveals that higher manual tracking adherence was associated the use of physical activity tracking devices, i.e. users who possessed a physical activity tracker (e.g. Fitbit, Apple Watch, Garmin) were more likely to complete the daily petal entries within the smartphone app.

There are several possible explanations for this correlation. First, those who are enthusiastic and engaged in tracking may be more likely to buy such a device, given their positive attitude and the value that they associate with it. Second, the practice of physical activity tracking (for which data collection is typically automated) may serve as a catalyst for engagement with the manual self-tracking. For example, users may have already established habits and routines with regards to wearing, charging and maintaining a device, around which self-reporting of data can be more easily integrated. Prior work has highlighted the importance of habit formation for long-term engagement [73]. Finally, as we discovered in Study 2, using a physical device can also serve as reminder to engage in the manual aspects of tracking. Users observe a salient reminder to record data each time they attend

to the physical device. While we cannot determine whether use of an accompanying tracking device, alongside a smartphone app for self-reporting health data, *causes* or *correlates* with greater adherence, our study suggests that clear reminders play an important role in adherent tracking.

7.6 Tracking Routine / Timing and Adherence

Tracking during the evening hours was the final significant predictor of participants' overall adherence, accounting for 11.5% of the variance. It is logical that participants tasked with reporting on a daily basis will be more inclined to enter values once they are able to recall and reflect on their health towards the end of the day. 74.3% of all entries took place during the evening hours (18:00 - 00:00).

However, 3.2% of all data entries occurred during the night (00:00 - 06:00), 11.0% occurred in the morning (06:00-12:00) and 11.5% in afternoon (12:00 - 18:00) on the day to which the data corresponded. Participants tracking at these times were less adherent. These data entries may account for participants pre-empting their symptoms for the rest of the day, for example if they are feeling confident that their symptoms are unlikely to change. This corresponds to our earlier finding that patients were less likely to track when their condition was stable or improving. Alternatively, earlier data entries may reflect patients who are responding to changes in their condition (e.g. sudden onset of pain) and recording data about these changes 'in the moment'. A reactive approach to tracking may lead to lower adherence when a condition is stable, compared with those who establish a tracking habit, such as recording data every evening, regardless of their health.

It appears that the time-based reminders within uMotif supported engagement with the app towards the end of the day. Similar findings have been reported in prior work on the effects of situational factors, including time of day [61], for gathering self-reported data. Prior work has also highlighted the need for contextual cues that are associated with locations [10, 64] and tasks (e.g. getting ready for bed) in order to better support habitual behaviours and the formation of new routines [74].

8 GENERALISABILITY OF FINDINGS

In this paper, we have focused on axSpA as an example of a chronic condition and attempted to uncover factors that influence adherence to daily self-reporting via a smartphone app. We anticipate that some of the predictors of adherence that we have identified are likely to generalise to other conditions. For example, many chronic musculoskeletal conditions are characterised by a similar group of symptoms, such as pain, fatigue, disturbed sleep and reduced mood, and affect a similarly broad age range. There are also common approaches to treatment and long-term management across such conditions. Furthermore, the fluctuating nature of axSpA, with cycles of flare versus remission, is also common in other chronic inflammatory conditions, such as rheumatoid arthritis, Crohn's disease, ulcerative colitis, and inflammatory bowel disease. Disease flares have been reported as a reason for increased app engagement both in the present study and prior studies of other conditions, e.g. [67].

However, generalising specific predictors of adherence to dissimilar conditions has clear limitations (e.g. where pain is not a common symptom). We posit that the general notion of condition severity is likely to remain an important factor that affects adherence, however the *specific* indicators of severity may change across conditions. For example, *fatigue* may be a more important determinant of tracking adherence in conditions such as ME/CFS.

Certain factors which we have discussed as being related to adherence (e.g. tracking hot flushes) may be relatively unique to axSpA. Equally, other conditions may have unique and useful prognostic data, which may increase adherence if they are able to be tracked.

The uMotif app studied in this paper has previously been used for the self-management of other diseases such as Parkinson's [40], and for research into symptoms in other rheumatic and musculoskeletal diseases [26]. It exhibits features common to many self-tracking apps beyond the rheumatic and musculoskeletal diseases self-management sphere – e.g. MONARCA [4, 30]. Hence we expect our findings to be relevant to many apps

that enable manual tracking of multiple symptoms and behaviours, however we note that issues of usability emerged as a clear determinant of adherence.

Beyond physical health conditions, self-tracking for mental wellness is becoming increasingly common [39]. The results of our work are less likely to generalise to acute mental health conditions. Prior work has shown that the precariousness and severity of such conditions can compromise adherence to self-tracking. For example, poor mental health can often be associated with experiences of “confusion ... and self-doubt” [36], leading users’ to distrust their own self-reported data and undermining their desire to track [48].

Prior work examining tracking in specific mental health conditions, such as Bipolar Disorder, indicates that there are important condition-specific and idiosyncratic factors that may impact adherence. For example, increased engagement in avoidance behaviours as a result of the condition can impact technology usage.

9 LIMITATIONS AND FUTURE WORK

We do not attempt to provide a comprehensive account of *all* determinants of adherence in this paper. Moreover, we contend that it is impossible to account for all possible variations in human behaviour and motivation relating to adherence. There are likely other factors (e.g. personality traits), which are not captured within our analyses. Our quantitative investigation was limited to exploration of data that was available through the app, and which had been collected for research purposes by the hospital.

Participants in Study 1 were initially encouraged by their clinician to enter data on a daily basis and to continue this for as long as possible. They were also told that their data may be used to inform management of their condition, as well as research about the condition more broadly. Anticipating that their data may be used in this way may have increased their adherence, but we are unable to determine the impact of these factors on overall adherence. Guidance from a clinician may have been a strong determinant of adherence for some and less important for others. Similarly, participants motivated by altruism may have stayed engaged even if they have little interest in self-tracking [7]. These factors should be explored further in future work, for example quantifying differences in adherence when data is not shared with clinicians.

Our analysis approach used a linear regression model to identify significant determinants of adherence. As highlighted by the qualitative finding relating to the non-monotonic relationship between pain and adherence, linear models may not accurately characterise all such relationships. Future work may explore the use of more complex models in order to understand adherence.

Finally, our longitudinal study focused on the use of one specific tracking app, *uMotif*. Although this app is generally well-designed and was reported as being easy to use by many participants, our survey results (which relate to a broader range of apps and technologies) indicate that user experience issues impact on adherence. Hence, different apps may be subject to varying rates of adherence on the basis of user experience factors. Future work could attempt to explore these factors further, for instance by measuring perceived ease of use (e.g. with SUS [12]) and investigating the extent to which this impacts daily self-tracking.

10 CONCLUSION

Digital technologies hold significant potential for improving the wellbeing of people living with chronic health conditions. The need for remote tracking of patients’ symptoms and behaviours is exacerbated by increased pressure on ‘in-person’ healthcare systems, for example during the COVID-19 pandemic. However, the ability of users to manage their condition through self-tracking relies on adherence with regular data collection. In this paper, we investigated factors that influence longitudinal adherence with smartphone-based self-tracking in a sample of axSpA patients. Our data suggests that adherence is impacted by entering data in the evening; the age of the user; selecting to track problematic symptoms; smartphone operating system; use of a physical activity

tracker; and mean average pain rating. User experience issues with devices may also affect adherence by making data entry difficult or affecting users' motivations to track.

Although the present study focused on a single chronic condition, axSpA, our findings have, to an extent, been echoed in the literature on other chronic, fluctuating conditions such as rheumatoid arthritis and chronic pain. We predict that similar patterns of adherence and engagement may be present in other chronic, inflammatory, fluctuating conditions of a similar nature. Furthermore, certain factors such as age, time/routine of tracking, and use of a physical activity tracker may be more widely applicable to app engagement within the general population. Future work on tracking app engagement should consider these factors.

11 ACKNOWLEDGEMENTS

This work is supported by funding from the Sir Halley Stewart Trust (Grant Reference: 2316). Any views expressed within this report are those of the authors and not necessarily those of the Trust. We gratefully acknowledge UCB for funding use of the uMotif application for the research studies, and the National Axial Spondyloarthritis Society and Bath Institute for Rheumatic Diseases for supporting patient and public engagement aspects of this research. We are grateful to Royal United Hospital, Bath for providing funding towards Violet Henderson's PhD studentship. We thank the members of the Project Nightingale team and the BathSPARC (Bath Spondyloarthritis Research Consortium) for their input and discussions relating to this research and all of the study participants for their valuable contributions. Due to confidentiality agreements with research collaborators, supporting data can only be made available to bona fide researchers subject to a data sharing agreement. Details of the self-tracking dataset are available at <https://www.projectnightingale.org/publications/>

REFERENCES

- [1] [n.d.]. Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. ([n. d.]).
- [2] Jessica S Ancker, Holly O Witteman, Baria Hafeez, Thierry Provencher, Mary Van de Graaf, and Esther Wei. 2015. "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), e202.
- [3] 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*. Association for Computing Machinery, New York, NY, USA, 6889–6901. <https://doi.org/10.1145/3025453.3025869>
- [4] Jakob E. Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. 2013. Designing Mobile Health Technology for Bipolar Disorder: A Field Trial of the Monarca System. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2627–2636. <https://doi.org/10.1145/2470654.2481364>
- [5] Rosie Barnett, Simon Jones, Matthew Young, and Raj Sengupta. 2019. Association of self-reported pain, fatigue, sleep, exercise, mood and stress in spondyloarthritis: initial analyses from the Project Nightingale study. In *BRITSpA 5th Annual Scientific Meeting*.
- [6] Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone, and Jeff Wilson. 2013. Health Mashups: Presenting Statistical Patterns between Wellbeing Data and Context in Natural Language to Promote Behavior Change. *ACM Trans. Comput.-Hum. Interact.* 20, 5, Article 30 (Nov. 2013), 27 pages. <https://doi.org/10.1145/2503823>
- [7] Anna L. Beukenhorst, Kelly Howells, Louise Cook, John McBeth, Terence W O'Neill, Matthew J Parkes, Caroline Sanders, Jamie C Sergeant, Katy S Weihrich, and William G Dixon. 2020. Engagement and Participant Experiences With Consumer Smartwatches for Health Research: Longitudinal, Observational Feasibility Study. *JMIR Mhealth Uhealth* 8, 1 (29 Jan 2020), e14368. <https://doi.org/10.2196/14368>
- [8] Pavol Bielik, Michal Tomlein, Peter Krátky, Štefan Mitřík, Michal Barla, and Mária Bieliková. 2012. Move2Play: An Innovative Approach to Encouraging People to Be More Physically Active. In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium (IHI '12)*. Association for Computing Machinery, New York, NY, USA, 61–70. <https://doi.org/10.1145/2110363.2110374>
- [9] Susanne Boll, Wilko Heuten, and Jochen Meyer. 2016. From Tracking to Personal Health. *Interactions* 23, 1 (Dec. 2016), 72–75. <https://doi.org/10.1145/2843586>
- [10] Mehdi Boukhechba, Lihua Cai, Philip I. Chow, Karl Fua, Matthew S. Gerber, Bethany A. Teachman, and Laura E. Barnes. 2018. Contextual Analysis to Understand Compliance with Smartphone-Based Ecological Momentary Assessment. In *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '18)*. Association for Computing Machinery, New York, NY, USA, 232–238. <https://doi.org/10.1145/3240925.3240967>

- [11] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597.
- [12] John Brooke. 2013. SUS: a retrospective. *Journal of Usability Studies* 8, 2 (2013), 29–40.
- [13] Clara Caldeira and Yunan Chen. 2019. Seniors and Self-tracking Technology. In *Perspectives on Human-Computer Interaction Research with Older People*, Sergio Sayago (Ed.). Springer International Publishing, Cham, 67–79. https://doi.org/10.1007/978-3-030-06076-3_5
- [14] Fallon E Chipidza, Rachel S Wallwork, and Theodore A Stern. 2015. Impact of the doctor-patient relationship. *The primary care companion for CNS disorders* 17, 5 (2015).
- [15] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. 2017. Understanding Self-Reflection: How People Reflect on Personal Data through Visual Data Exploration. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '17)*. Association for Computing Machinery, New York, NY, USA, 173–182. <https://doi.org/10.1145/3154862.3154881>
- [16] Eun Kyoung Choe, Nicole B. Lee, Bongshin Lee, Wanda Pratt, and Julie A. Kientz. 2014. Understanding Quantified-Selfers' Practices in Collecting and Exploring Personal Data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1143–1152. <https://doi.org/10.1145/2556288.2557372>
- [17] Chia-Fang Chung, Jonathan Cook, Elizabeth Bales, Jasmine Zia, and Sean A Munson. 2015. 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 (21 Aug 2015), e203. <https://doi.org/10.2196/jmir.4364>
- [18] James Clawson, Jessica A Pater, Andrew D Miller, Elizabeth D Mynatt, and Lena Mamykina. 2015. No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. *Proceedings of the 2015 International Conference on Ubiquitous Computing (2015)*, 647–658. <http://dx.doi.org/10.1145/2750858.2807554>
- [19] Sunny Consolvo, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, Ian Smith, and James A. Landay. 2008. Activity Sensing in the Wild: A Field Trial of Ubift Garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. Association for Computing Machinery, New York, NY, USA, 1797–1806. <https://doi.org/10.1145/1357054.1357335>
- [20] Felicia Cordeiro, Daniel A. Epstein, Edison Thomaz, Elizabeth Bales, Arvind K. Jagannathan, Gregory D. Abowd, and James Fogarty. 2015. Barriers and Negative Nudges: Exploring Challenges in Food Journaling. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1159–1162. <https://doi.org/10.1145/2702123.2702155>
- [21] Andrea Cuttone, Michael Kai Petersen, and Jakob Eg Larsen. 2014. Four Data Visualization Heuristics to Facilitate Reflection in Personal Informatics. In *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice*, Constantine Stephanidis and Margherita Antona (Eds.). Springer International Publishing, Cham, 541–552.
- [22] Tabby Davies, Simon L. Jones, and Ryan M. Kelly. 2019. Patient Perspectives on Self-Management Technologies for Chronic Fatigue Syndrome. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300452>
- [23] Aiden Doherty, Dan Jackson, Nils Hammerla, Thomas Plötz, Patrick Olivier, Malcolm H Granat, Tom White, Vincent T Van Hees, Michael I Trenell, Christopher G Owen, et al. 2017. Large scale population assessment of physical activity using wrist worn accelerometers: The UK Biobank Study. *PLoS one* 12, 2 (2017), e0169649.
- [24] Angelika Dohr, Jeff Engler, Frank Bentley, and Richard Whalley. 2012. Gluballoon: An Unobtrusive and Educational Way to Better Understand One's Diabetes. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. Association for Computing Machinery, New York, NY, USA, 665–666. <https://doi.org/10.1145/2370216.2370357>
- [25] Adam Drewnowski, Elaine Monsen, Diana Birkett, Susan Gunther, Susan Vendeland, Jeff Su, and Gillian Marshall. 2003. Health screening and health promotion programs for the elderly. *Disease Management & Health Outcomes* 11, 5 (2003), 299–309.
- [26] Katie L Druce, David S Gibson, Keven McEleney, Stephanie Meleck, Ben James, Bruce Hellman, William G Dixon, and John Mcbeth. 2020. P60Daily fluctuations in fatigue severity and C-reactive protein in people with rheumatic diseases. *Rheumatology* 59, 2 (04 2020). <https://doi.org/10.1093/rheumatology/keaa111.059>
- [27] Katie L Druce, John McBeth, Sabine N van der Veer, David A Selby, Bertie Vidgen, Konstantinos Georgatzis, Bruce Hellman, Rashmi Lakshminarayana, Afiqu Chowdhury, David M Schultz, Caroline Sanders, Jamie C Sergeant, and William G Dixon. 2017. Recruitment and Ongoing Engagement in a UK Smartphone Study Examining the Association Between Weather and Pain: Cohort Study. *JMIR Mhealth Uhealth* 5, 11 (01 Nov 2017), e168. <https://doi.org/10.2196/mhealth.8162>
- [28] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A Lived Informatics Model of Personal Informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 731–742. <https://doi.org/10.1145/2750858.2804250>
- [29] Ahmed Fadhil. 2019. Beyond technical motives: Perceived user behavior in abandoning wearable health & wellness trackers. *arXiv preprint arXiv:1904.07986* (2019).

- [30] Mads Frost, Afsaneh Doryab, Maria Faurholt-Jepsen, Lars Vedel Kessing, and Jakob E. Bardram. 2013. Supporting Disease Insight through Data Analysis: Refinements of the Monarca Self-Assessment System. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. Association for Computing Machinery, New York, NY, USA, 133–142. <https://doi.org/10.1145/2493432.2493507>
- [31] Anda Gershon, Christopher N Kaufmann, John Torous, Colin Depp, and Terence A Ketter. 2019. Electronic Ecological Momentary Assessment (EMA) in youth with bipolar disorder: Demographic and clinical predictors of electronic EMA adherence. *Journal of Psychiatric Research* 116 (2019), 14–18.
- [32] Friedrich M Götz, Stefan Stieger, and Ulf-Dietrich Reips. 2017. Users of the main smartphone operating systems (iOS, Android) differ only little in personality. *PloS one* 12, 5 (2017), e0176921.
- [33] Vicki L Hanson. 2010. Influencing technology adoption by older adults. *Interacting with Computers* 22, 6 (2010), 502–509.
- [34] Gabriella M. Harari, Weichen Wang, Sandrine R. Müller, Rui Wang, and Andrew T. Campbell. 2017. Participants' Compliance and Experiences with Self-Tracking Using a Smartphone Sensing App. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (UbiComp '17)*. Association for Computing Machinery, New York, NY, USA, 57–60. <https://doi.org/10.1145/3123024.3123164>
- [35] William Hue, Simon Jones, and Raj Sengupta. 2019. Exploring the Future Role of Self-Tracking Data in the Rheumatology Clinic. *Studies in health technology and informatics* 259 (2019), 33–38.
- [36] Maree L Inder, Marie T Crowe, Peter R Joyce, Stephanie Moor, Janet D Carter, and Sue E Luty. 2010. “I really don't know whether it is still there”: Ambivalent acceptance of a diagnosis of bipolar disorder. *Psychiatric Quarterly* 81, 2 (2010), 157–165.
- [37] Mirou Jaana and Guy Pare. 2020. Use of Mobile Health Technologies for Self-tracking Purposes among Seniors: A Comparison to the General Adult Population in Canada. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*.
- [38] Ravi Karkar, Jessica Schroeder, Daniel A. Epstein, Laura R. Pina, Jeffrey Scofield, James Fogarty, Julie A. Kientz, Sean A. Munson, Roger Vilardaga, and Jasmine Zia. 2017. TummyTrials: A Feasibility Study of Using Self-Experimentation to Detect Individualized Food Triggers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 6850–6863. <https://doi.org/10.1145/3025453.3025480>
- [39] Christina Kelley, Bongshin Lee, and Lauren Wilcox. 2017. Self-Tracking for Mental Wellness: Understanding Expert Perspectives and Student Experiences. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 629–641. <https://doi.org/10.1145/3025453.3025750>
- [40] Rashmi Lakshminarayana, Duolao Wang, David Burn, K Ray Chaudhuri, Clare Galtrey, Natalie Valle Guzman, Bruce Hellman, Ben James, Suvankar Pal, Jon Stamford, et al. 2017. Using a smartphone-based self-management platform to support medication adherence and clinical consultation in Parkinson's disease. *NPJ Parkinson's disease* 3, 1 (2017), 1–10.
- [41] J. Richard Landis and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (1977), 159–174. <http://www.jstor.org/stable/2529310>
- [42] Amanda Lazar, Christian Koehler, Theresa Jean Tanenbaum, and David H. Nguyen. 2015. Why We Use and Abandon Smart Devices. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 635–646. <https://doi.org/10.1145/2750858.2804288>
- [43] Dan Ledger and Daniel McCaffrey. 2014. Inside wearables: How the science of human behavior change offers the secret to long-term engagement. *Endeavour Partners* 200, 93 (2014), 1.
- [44] Elaine A Leventhal, Howard Leventhal, Pamela Schaefer, and Douglas Easterling. 1993. Conservation of energy, uncertainty reduction, and swift utilization of medical care among the elderly. *Journal of Gerontology* 48, 2 (1993), P78–P86.
- [45] Ian Li, Anind K. Dey, and Jodi Forlizzi. 2011. Understanding My Data, Myself: Supporting Self-Reflection with UbiComp Technologies. In *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp '11)*. Association for Computing Machinery, New York, NY, USA, 405–414. <https://doi.org/10.1145/2030112.2030166>
- [46] Christopher Steven Marcum. 2013. Age differences in daily social activities. *Research on aging* 35, 5 (2013), 612–640.
- [47] Tania S Marin, Constance Kourbelis, Jonathon Foote, Peter Newman, Alex Brown, Mark Daniel, Neil T Coffee, Stephen J Nicholls, Anand Ganesan, Vincent L Versace, et al. 2019. Examining adherence to activity monitoring devices to improve physical activity in adults with cardiovascular disease: A systematic review. *European Journal of Preventive Cardiology* 26, 4 (2019), 382–397.
- [48] Mark Matthews, Elizabeth Murnane, and Jaime Snyder. 2017. Quantifying the Changeable Self: The role of self-tracking in coming to terms with and managing bipolar disorder. *Human-Computer Interaction* 32, 5-6 (2017), 413–446.
- [49] Mollie McKillop, Lena Mamykina, and Noémie Elhadad. 2018. Designing in the Dark: Eliciting Self-Tracking Dimensions for Understanding Enigmatic Disease. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3173574.3174139>
- [50] Siobhan K McMahon, Beth Lewis, Michael Oakes, Weihua Guan, Jean F Wyman, and Alexander J Rothman. 2016. Older adults' experiences using a commercially available monitor to self-track their physical activity. *JMIR mHealth and uHealth* 4, 2 (2016), e35.
- [51] Helena M. Mentis, Anita Komlodi, Katrina Schrader, Michael Phipps, Ann Gruber-Baldini, Karen Yarbrough, and Lisa Shulman. 2017. Crafting a View of Self-Tracking Data in the Clinical Visit. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing*

- Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5800–5812. <https://doi.org/10.1145/3025453.3025589>
- [52] Jochen Meyer, Merlin Wasmann, Wilko Heuten, Abdallah El Ali, and Susanne C.J. Boll. 2017. Identification and Classification of Usage Patterns in Long-Term Activity Tracking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 667–678. <https://doi.org/10.1145/3025453.3025690>
- [53] Jun-Ki Min, Afsaneh Doryab, Jason Wiese, Shahriyar Amini, John Zimmerman, and Jason I. Hong. 2014. Toss 'n' Turn: Smartphone as Sleep and Sleep Quality Detector. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 477–486. <https://doi.org/10.1145/2556288.2557220>
- [54] Sonali R. Mishra, Predrag Klasnja, John MacDuffie Woodburn, Eric B. Hekler, Larsson Omberg, Michael Kellen, and Lara Mangravite. 2019. Supporting Coping with Parkinson's Disease Through Self Tracking. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–16. <https://doi.org/10.1145/3290605.3300337>
- [55] Elizabeth L. Murnane, Xin Jiang, Anna Kong, Michelle Park, Weili Shi, Connor Soohoo, Luke Vink, Iris Xia, Xin Yu, John Yang-Sammataro, Grace Young, Jenny Zhi, Paula Moya, and James A. Landay. 2020. Designing Ambient Narrative-Based Interfaces to Reflect and Motivate Physical Activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376478>
- [56] B Mitchell Peck. 2011. Age-related differences in doctor-patient interaction and patient satisfaction. *Current Gerontology and Geriatrics Research* 2011 (2011).
- [57] Frederic W Platt, David L Gaspar, John L Coulehan, Lucy Fox, Andrew J Adler, W Wayne Weston, Robert C Smith, and Moira Stewart. 2001. "Tell me about yourself": the patient-centered interview. *Annals of Internal Medicine* 134, 11 (2001), 1079–1085.
- [58] Rüdiger Pryss, Manfred Reichert, Winfried Schlee, Myra Spiliopoulou, Berthold Langguth, and Thomas Probst. 2018. Differences between android and ios users of the trackyourtinnitus mobile crowdsensing mhealth platform. In *2018 IEEE 31st international symposium on computer-based medical systems (CBMS)*. IEEE, 411–416. <https://doi.org/10.1109/CBMS.2018.00078>
- [59] Tom Quisel, Luca Foschini, Susan M Zbikowski, and Jessie L Juusola. 2019. The association between medication adherence for chronic conditions and digital health activity tracking: retrospective analysis. *Journal of Medical Internet Research* 21, 3 (2019), e11486.
- [60] Md Mahbubur Rahman, Viswam Nathan, Ebrahim Nemati, Korosh Vatanparvar, Mohsin Ahmed, and Jilong Kuang. 2019. Towards Reliable Data Collection and Annotation to Extract Pulmonary Digital Biomarkers Using Mobile Sensors. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth'19)*. Association for Computing Machinery, New York, NY, USA, 179–188. <https://doi.org/10.1145/3329189.3329204>
- [61] Alex T Ramsey, Julie Loebach Wetherell, Colin Depp, David Dixon, and Eric Lenze. 2016. Feasibility and acceptability of smartphone assessment in older adults with cognitive and emotional difficulties. *Journal of Technology in Human Services* 34, 2 (2016), 209–223.
- [62] Samuel Reade, Karen Spencer, Jamie C Sergeant, Matthew Sperrin, David M Schultz, John Ainsworth, Rashmi Lakshminarayana, Bruce Hellman, Ben James, John McBeth, et al. 2017. Cloudy with a chance of pain: engagement and subsequent attrition of daily data entry in a smartphone pilot study tracking weather, disease severity, and physical activity in patients with rheumatoid arthritis. *JMIR mHealth and uHealth* 5, 3 (2017), e37.
- [63] Aki Rintala, Martien Wampers, Inez Myin-Germeys, and Wolfgang Viechtbauer. 2019. Response compliance and predictors thereof in studies using the experience sampling method. *Psychological Assessment* 31, 2 (2019), 226.
- [64] Aki Rintala, Martien Wampers, Inez Myin-Germeys, and Wolfgang Viechtbauer. 2020. Momentary predictors of compliance in studies using the experience sampling method. *Psychiatry Research* (2020), 112896.
- [65] James P Sanders, Adam Loveday, Natalie Pearson, Charlotte Edwardson, Thomas Yates, Stuart JH Biddle, and Dale W Esliger. 2016. Devices for self-monitoring sedentary time or physical activity: a scoping review. *Journal of Medical Internet Research* 18, 5 (2016), e90.
- [66] Jessica Schroeder, Chia-Fang Chung, Daniel A. Epstein, Ravi Karkar, Adele Parsons, Natalia Murinova, James Fogarty, and Sean A. Munson. 2018. Examining Self-Tracking by People with Migraine: Goals, Needs, and Opportunities in a Chronic Health Condition. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 135–148. <https://doi.org/10.1145/3196709.3196738>
- [67] Bart F Seppen, Jimmy Wiegel, Merel J L'ami, Sharon Duarte dos Santos Rico, Fabio S Catarinella, Franktien Turkstra, Maarten Boers, and Wouter H Bos. 2020. Feasibility of Self-Monitoring Rheumatoid Arthritis With a Smartphone App: Results of Two Mixed-Methods Pilot Studies. *JMIR Form Res* 4, 9 (21 Sep 2020), e20165. <https://doi.org/10.2196/20165>
- [68] Heather Shaw, David A Ellis, Libby-Rae Kendrick, Fenja Ziegler, and Richard Wiseman. 2016. Predicting smartphone operating system from personality and individual differences. *Cyberpsychology, Behavior, and Social Networking* 19, 12 (2016), 727–732.
- [69] Patrick C Shih, Kyungsik Han, Erika Shehan Poole, Mary Beth Rosson, and John M Carroll. 2015. Use and adoption challenges of wearable activity trackers. *ICConference 2015 Proceedings* (2015).
- [70] Stefan Siebert, Raj Sengupta, and Alexander Tsoukas. 2016. *Axial spondyloarthritis*. Oxford University Press.
- [71] Joachim Sieper and Denis Poddubnyy. 2017. Axial spondyloarthritis. *The Lancet* 390, 10089 (2017), 73–84.
- [72] Alexander W Sokolovsky, Robin J Mermelstein, and Donald Hedeker. 2014. Factors predicting compliance to ecological momentary assessment among adolescent smokers. *Nicotine & Tobacco Research* 16, 3 (2014), 351–358.

- [73] Katarzyna Stawarz, Anna L. Cox, and Ann Blandford. 2015. Beyond Self-Tracking and Reminders: Designing Smartphone Apps That Support Habit Formation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2653–2662. <https://doi.org/10.1145/2702123.2702230>
- [74] Katarzyna Stawarz, Marcela D Rodríguez, Anna L Cox, and Ann Blandford. 2016. Understanding the use of contextual cues: design implications for medication adherence technologies that support remembering. *Digital health* 2 (2016), 2055207616678707.
- [75] Vibeke Strand and Jasvinder A Singh. 2017. Patient burden of axial spondyloarthritis. *Journal of Clinical Rheumatology* 23, 7 (2017), 383.
- [76] Dallas Swendeman, Nithya Ramanathan, Laura Baetscher, Melissa Medich, Aaron Scheffler, W Scott Comulada, and Deborah Estrin. 2015. Smartphone self-monitoring to support self-management among people living with HIV: perceived benefits and theory of change from a mixed-methods, randomized pilot study. *Journal of Acquired Immune Deficiency Syndromes (1999)* 69, 0 1 (2015), S80.
- [77] Lie Ming Tang, Margot Day, Lina Engelen, Philip Poronnik, Adrian Bauman, and Judy Kay. 2016. Daily & Hourly Adherence: Towards Understanding Activity Tracker Accuracy. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 3211–3218. <https://doi.org/10.1145/2851581.2892438>
- [78] Lie Ming Tang, Jochen Meyer, Daniel A. Epstein, Kevin Bragg, Lina Engelen, Adrian Bauman, and Judy Kay. 2018. Defining Adherence: Making Sense of Physical Activity Tracker Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 37 (March 2018), 22 pages. <https://doi.org/10.1145/3191769>
- [79] David H Thom, Kurt M Ribisl, Anita L Stewart, Douglas A Luke, and The Stanford Trust Study Physicians. 1999. Further validation and reliability testing of the Trust in Physician Scale. *Medical Care* (1999), 510–517.
- [80] John Torous, Rohn Friedman, and Matcheri Keshavan. 2014. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. *JMIR mHealth and uHealth* 2, 1 (2014), e2.
- [81] Tiffany Tran and Dan Nathan-Roberts. 2018. Design Considerations in Wearable Technology for Patients with Bipolar Disorder. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 62, 1, 1187–1191. <https://doi.org/10.1177/1541931218621273> arXiv:<https://doi.org/10.1177/1541931218621273>
- [82] Stewart G Trost, Kerry L Mciver, and Russell R Pate. 2005. Conducting accelerometer-based activity assessments in field-based research. *Medicine and Science in Sports and Exercise* 37, 11 Suppl (2005), S531–43.
- [83] Gabrielle M Turner-McGrievy, Caroline Glagola Dunn, Sara Wilcox, Alycia K Boutté, Brent Hutto, Adam Hoover, and Eric Muth. 2019. Defining adherence to mobile dietary self-monitoring and assessing tracking over time: tracking at least two eating occasions per day is best marker of adherence within two different mobile health randomized weight loss interventions. *Journal of the Academy of Nutrition and Dietetics* 119, 9 (2019), 1516–1524.
- [84] Niels van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The Experience Sampling Method on Mobile Devices. *ACM Comput. Surv.* 50, 6, Article 93 (Dec. 2017), 40 pages. <https://doi.org/10.1145/3123988>
- [85] Edward Jay Wang, Junyi Zhu, Mohit Jain, Tien-Jui Lee, Elliot Saba, Lama Nachman, and Shwetak N. Patel. 2018. Seismo: Blood Pressure Monitoring Using Built-in Smartphone Accelerometer and Camera. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3173574.3173999>
- [86] Dianne S Ward, Kelly R Evenson, Amber Vaughn, Anne Brown Rodgers, and Richard P Troiano. 2005. Accelerometer use in physical activity: best practices and research recommendations. *Medicine & Science in Sports & Exercise* 37, 11 (2005), S582–S588.
- [87] Peter West, Max Van Kleek, Richard Giordano, Mark Weal, and Nigel Shadbolt. 2017. Information quality challenges of patient-generated data in clinical practice. *Frontiers in Public Health* 5 (2017), 284.
- [88] Xiaomeng Xu, Samantha Tupy, Stephen Robertson, Ashley L Miller, Danielle Correll, Rick Tivis, and Claudio R Nigg. 2018. Successful adherence and retention to daily monitoring of physical activity: Lessons learned. *PloS one* 13, 9 (2018), e0199838.
- [89] Leah R Yingling, Valerie Mitchell, Colby R Ayers, Marlene Peters-Lawrence, Gwenyth R Wallen, Alyssa T Brooks, James F Troendle, Joel Adu-Brimpong, Samantha Thomas, JaWanna Henry, et al. [n.d.]. Adherence with physical activity monitoring wearable devices in a community-based population: observations from the Washington, DC, Cardiovascular Health and Needs Assessment. *Translational Behavioral Medicine* 7, 4 ([n. d.]).