Cloudy with a Chance of Pain:

an interactive study investigating the link between

symptoms of chronic pain and the weather

Word count: 9,983 (excluding references)

Declaration of Originality
I confirm that this report is my own work. I have: • Read and understood the plagiarism guidelines • Clearly referenced all sources used in the work • Fully referenced and used inverted commas for all text quoted from journals, web etc. • Provided the sources for all tables, figures, data etc. that are not my own work • Acknowledged the work of any other(s) that is not my own
I am aware of and understand the University's policy on plagiarism and I certify that this assignment is my own work, except where indicated by referencing.

ABSTRACT

BACKGROUND

Cloudy with a Chance of Pain is a large mHealth study which aims to examine the putative association between chronic pain and the weather. Characteristics and beliefs of the study population will be described and the degree of participant engagement will be investigated and quantified.

METHODS

Baseline descriptive analyses of the study population (n=11,042) were carried out. Dates of participant enrolment, how often they entered data, and how long they stayed in the study were used to assess engagement. Hidden Markov models divided participants into clusters depending on their engagement. Answers to feedback questionnaires were analysed to determine factors contributing to user engagement and disengagement.

RESULTS

The majority of participants were female (80.35%), the mean age was 49 and the most common diagnoses were arthritis and fibromyalgia. 4 clusters of engagement were described: high, medium, low and “tourist”. Those in the high engagement cluster were significantly older. 15.76% of participants stayed engaged for 6 months or more. The main factors contributing to patient disengagement were issues with the app, lifestyle factors, amount of time required to participate and technical difficulties.

CONCLUSION

Cloudy has demonstrated that mHealth apps are a feasible way of collecting extensive amounts of data from large study populations over extended periods of time. There are varying degrees of participant engagement with such apps due to multiple factors such as age, beliefs and personal preferences. It is important that mHealth apps are appealing and accessible to target study populations in order to maximise user engagement.

1. INTRODUCTION

1.1 Chronic pain and the weather

Chronic pain is a common problem which is estimated to affect almost one-fifth of the European population[1]. Non-malignant chronic pain is traditionally defined as pain that persists for at least 3 months[2]. It has a considerable impact on patients' quality of life[1] and is associated with many demographic and social factors including female gender, increasing age, employment status, geographical situation and cultural beliefs[3].

An association between the weather and chronic disease has been speculated since the time of Hippocrates in 400 B.C[4]. Since then, there have been numerous investigations which have resulted in inconclusive evidence that the weather has an impact on the severity of chronic pain. Most studies to date have focussed on rheumatoid arthritis[5,6] and osteoarthritis[7,8,9]; however, the effect of weather on fibromyalgia[10,11], lower back pain[12] and other long-term problems such as temporomandibular pain[13] has also been explored.

Most previous studies have used patient-reported pain scores and retrospectively matched these with weather data from local or national meteorological stations[6,12]. Although, there have been some studies which have either recorded weather data in real-time by using information from meteorological institutes[14] or used portable data loggers to sample and record specific weather variables in the moment pain scores are recorded by the participant[13].

The interpretation of some prior studies is limited by their relatively small sample sizes and short periods of data collection[6,13,14]. Moreover, an issue with using data from weather stations local to participants' addresses is that they are assumed to consistently be in the same location, whilst this may not always be the case[12]. A potential way of overcoming this would be to capture weather data for a participant's exact location using Global Positioning System (GPS) – whereby the exact co-ordinates of a device's location are recorded and subsequently linked to weather data[15]. Recruitment of more participants to increase study reliability, along with longitudinally recording reliable weather data and participant-reported symptoms in real time, could be achieved by taking advantage of mobile health technologies.

1.2 mHealth technologies

Mobile health (mHealth) is defined by the World Health Organization's (WHO) Global Observatory for eHealth (GOe) as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices"[16]. There have been significant developments in the field of mHealth both in the delivery and collection of healthcare data[17]. It offers benefits over traditional approaches to data collection – namely face-to-face and pen-and-paper methods – in that it has the potential to reduce the financial cost, time and manpower required to successfully obtain a large amount of information[18]. Smartphones offer a unique opportunity to collect population data on a substantial scale because they are already so widely used. It is estimated that there are over 2 billion smartphone users worldwide and this number is only set to increase[19]. A benefit that smartphone applications (apps) offer is that they provide the facility to input data when it is required or convenient and they allow data to be collected and analysed in real time[20]. Smartphone apps have the capacity to integrate information inputted by users with GPS data which can then be linked to local weather stations. For the purpose of investigating the link between chronic pain and the weather, this is an exciting and feasible way to investigate relationships between worsening pain symptoms and specific weather variables which may be associated with them.

Patient-reported outcomes (PROs) are a useful way of monitoring the status of an individual's health condition and they allow large amounts of data to be collected remotely and over extended periods of time[21]. PROs can be collected using smartphone apps and this was done successfully in 2015 by the mPower study[22]. This large observational smartphone-based study required patients with a diagnosis of Parkinson's disease – a chronic neurological condition – to record their symptoms and response to medication every day. Similarly, information relating to sleep patterns and mood status was collected on a daily basis from patients with breast cancer who were receiving chemotherapy via a smartphone app in a 90-day longitudinal study[23]. Both of these studies have indicated that it is feasible to collect longitudinal data with multiple variables by means of a smartphone app.

1.3 Engagement with mHealth programmes

Interactive mHealth apps allow for a dynamic engagement between health care providers and patients[17]. However, this means that there can be varying degrees of patient participation – and there are still barriers to engaging patients with this method of data collection and maintaining their interest for extended periods[21,24,25]. Different approaches have been taken to measure engagement in mHealth studies. Examples include total data entry over an entire study period[26,27], individual data reporting over specified time intervals[23,28] and the number of days of participation in a study[29]. With regards to retaining participants over time, a randomised controlled trial investigating a weight loss programme[27] found that participant retention was 93% at 6 months with smartphones compared to 55% with a website and 53% with a diary. In addition, participant adherence was significantly higher with smartphones. This highlights that smartphones are an effective way of enhancing participant engagement over an extended period.

There are a variety of factors which are believed to influence the likelihood of consistent engagement with mHealth programmes. Such factors include participant characteristics, information provided before taking part in the study, accessibility of the study, ease of application use, reminders, workload involved, time required and available support[30]. In contrast, it has been observed that a substantial number of participants often drop out or do not fully carry out what is expected of them before completion of a mHealth trial. This trend has been termed by Eysenbach[30] as the "law of attrition" – whereby participants either stop using mHealth applications, are lost to follow-up, or both. Factors such as age, education level and employment status may hinder participant engagement[31]. Reasons for different degrees of engagement are summarised in Table 1[30,31].

Table 1: Factors influencing participant engagement and disengagement in mHealth studies

mHealth study factors	mHealth application factors	Participant factors
Information provided	Ease of use	Perceived benefits from taking part
Available support	Interface appearance	External events
Ease of enrolment	Reminders	Competing interests
Ease of discontinuing	Time required to complete	Demographics (age, gender, education, employment, socioeconomic factors)
	Amount of work involved	

In addition to the challenge of maintaining participant engagement, several mHealth studies have highlighted that varying levels of participant engagement exist within the groups which stay involved throughout[23,25,32]. In 2004, Anhoj *et al*[32] investigated the feasibility of diary data collection from asthma patients using mobile phone SMS messages. The response rates to SMS messages were analysed and the investigators split the participants into two groups: high compliers and low compliers. This study suggested that further mHealth investigations should involve larger sample sizes which were more representative of the overall patient population and that user segmentation – that is, different groups of participants who engage differently – is something to bear in mind. More recently, Min *et al*[23] calculated the rates of smartphone-app data entry for individual patients during a study to examine the relationship between sleep disturbance and breast cancer chemotherapy. The rate of patient-level reporting was distributed in a bi-modal fashion; hence the authors described the population in a similar manner by designating participants into high compliance and low compliance groups. It was concluded that the only variable which influenced the rate of compliance with the app was occupational status – with unemployed participants more likely to input data on a regular basis.

The most recent mHealth study to explore participant engagement was the Asthma Mobile Health Study[33]. This large-scale clinical observational study used Apple's ResearchKit to design an app which allowed participants to record the severity of their asthma symptoms and then correlate this to environmental factors including heat, pollen and wildfires. The investigators did attempt to quantify participant engagement however, although they identified different cohorts of participants with regards to their engagement based upon a minimum number of completed surveys, not all participants were selected to be designated to a cohort which corresponded to their level of participation. Importantly, it was acknowledged by the authors that selection bias and retaining participants are significant problems facing smartphone-app studies.

These studies illustrate that mHealth studies are challenged by differing levels of participant engagement. However there have been no clear reasons or evidence provided to suggest why this is so. Furthermore, there is a limited amount of literature describing the patterns of engagement in mHealth studies. Many groups have alluded to this phenomenon but have not explored it any further[24]. In the studies mentioned above, participants have been grouped into just two categories – of either high or low engagement – but this does not offer very meaningful information.

1.4 Cloudy with a Chance of Pain

Cloudy with a Chance of Pain is a prospective cohort study which aims to investigate relationships between perceived pain and the weather in patients suffering from chronic pain. By means of a smartphone app, "Cloudy", patient-reported pain symptoms were linked to geographical weather data to correlate pain severity to weather conditions. Due to its sizeable population and lengthy data collection period, Cloudy offers an opportunity to further explore the patterns of participant engagement in mHealth studies.

1.5 Aim of APEP

This report will describe the characteristics of the Cloudy study population and analyse participants' beliefs with regards to weather affecting their pain. The degrees of participant engagement will be investigated for the entire study population and further analysed to evaluate which factors might increase the likelihood of continuous engagement with a mHealth study. Previous work done to investigate Cloudy engagement[34,35] will be adapted and repeated in order to achieve this. Furthermore, Cloudy feedback questionnaires will be assessed to gauge which factors in particular contributed to participant engagement. Hopefully, by grouping participants into engagement groups and analysing their opinions of a large-scale smartphone study, guidance can be offered on how to best engage participants from target populations in future mHealth studies.

2. METHODS

University ethics approval was granted for Cloudy with a Chance of Pain (Ref:UREC4/15522).

2.1 Study population

Patients aged 17 years or older who were suffering from chronic pain, resident in the UK and owned a smartphone were eligible to take part in Cloudy with a Chance of Pain. Recruitment began in January 2016 and ended in January 2017. In order to recruit as many eligible participants as possible, the study was advertised through television, radio and newspaper. Participants required a smartphone with an iOS or Android operating system so they could download the free uMotif "Cloudy" app.

2.2 Baseline Data

Basic demographic data was gathered from participants using a baseline questionnaire. Users were prompted to complete this before entering any other data into the application. For each participant, the information requested included: gender, year of birth and the outward code from their postcode. Sites where pain was experienced were recorded by selecting one or more body parts (head, face, back etc.) from a given list or by choosing the option for pain at multiple sites or pain all over the body. In addition, participants recorded their existing beliefs about the relationship between pain and the weather using a series of multiple choice questions with answers on a scale ranging from 0 to 10 – with 0 corresponding to no belief of the association and 10 reflecting a strong belief in an association.

2.3 Inputting symptom data and collecting GPS weather data

Participants were encouraged to enter data into the application at least once a day for 6 months, or longer if they wished to do so. The Cloudy app featured a 10-segment motif where participants could enter a score (1-5) to reflect their pain severity, sleep quality, activity levels, mood and more. The motif is shown in Figure 1.

Figure 1: Cloudy app 10-segment motif

Figure 1: Screenshot of Cloudy symptom tracker with 10-segment motif. Participants clicked on a segment to enter a score of 1-5 to reflect pain severity, morning stiffness, tiredness on waking, level of fatigue, time spent outside, activity levels, disease severity, effect of weather that day on pain severity, wellbeing, and mood.

A reminder was sent at 6:42pm each day in the form of a "push notification" to prompt participants to enter data. GPS data from the participants' phone was passively requested on an hourly basis by the app. Location data was recorded once a day by iOS smartphones and multiple times a day for Android smartphones. This was then linked to hourly weather data from the nearest Met Office station. Data was collected until noon on 20/04/17.

2.4 Descriptive analysis of the study population

Everyone who downloaded the Cloudy app received a unique user identification number (user I.D.). Users who did not submit any symptom data over the entire study period were excluded. From this sample, only users who had submitted a baseline questionnaire were selected.

The baseline questionnaire data was analysed in Excel and RStudio. Participants who did not provide a year of birth, or had provided a year of birth that was impossible, had their age recorded as N/A. The average ages were calculated including the N/A values.

In order to find out where participants experienced pain (single site, multiple sites or all over the body) both Microsoft Excel and RStudio were used. Those who experienced pain all over, and those who had selected pain at multiple sites, were removed in order to determine how many participants experienced pain at a single site. As many users selected more than one pain site option but did not select pain at multiple sites or pain all over, participants with more than one site selected were categorised as having pain at multiple sites.

In the weather belief section, answers to a free-text question which asked participants if they thought anything else affected pain were analysed and broken down into other, unsure and no relationship.

2.5 Participant enrolment

Using Excel, the dates on which participants enrolled were ordered chronologically. The number of participants enrolling on each date was recorded in order to determine the principal time-points of high enrolment and assess the patterns of enrolment over the entire enrolment period. Answers to a multiple-choice question asking participants how they found out about the study were explored.

2.6 Engagement patterns

The date on which participants enrolled in the study and first entered data (F), their last day of data entry (L), the number of days they entered motif data for (n), and the study end date (E) were noted. The period of time over which a participant interacted with the app (PP) was determined by calculating the number of days between the day of their first (F) and last (L) data entries (L-F=PP). The total period over which participants were able to enter data (OP) was calculated as the time between the day they first enrolled (F) and entered data, and the final day (E) of data collection (E-F=OP). These figures were then used to determine each individual's percentage contribution of data entry. This was done both over the period which they entered data for (PP) and for the entire period that they could have entered data (OP). This was because some individuals submitted data regularly over a short period of time – which was reflected in a high percentage contribution of data entry – when in reality this was a small contribution over the total amount of days they could have potentially entered data.

The number of days (n) an individual entered data for was divided by their PP in order to find their personal percentage contribution. The number of days (n) they entered data for was divided by their OP to find their overall percentage contribution. This is shown below.

$$\frac{\text{number of days data entered (n)}}{\text{PP}} = \text{personal percentage contribution (\%)}$$

$$\frac{\text{number of days data entered (n)}}{\text{OP}} = \text{overall percentage contribution (\%)}$$

Initially, participants were asked to stay in the study for 6 months. In order to assess how many participants achieved this, 6 months was added to their start date (F) and any individuals who had an end date (E) which reached or surpassed this date were recorded. The formula used is shown below. Participants who enrolled less than 6 months before the study end-date (20/10/16) were not included in the analysis.

contributed 6 months or more: $E - (F + 6 \text{ months}) \geq 0$

Previous work by Selby[34] and Druce *et al*[35] used a Hidden Markov Model (HMM) to measure participant engagement. The HMM allowed three states of user engagement to be defined: high, low and disengaged. Participants could interchange between these states during the study period. Subsequently, users were clustered according to their likelihood of transitioning between states. This culminated in 4 clusters: active, moderate, inactive and tourist.

The process was adapted and re-run for this larger study population in order to assign each participant to a "cluster" corresponding to a degree of engagement. The HMM work-up and clustering was carried out by a PhD student prior to the analysis done in this report. The characteristics of each cluster were then investigated and compared by age, sex and level of engagement.

2.7 Feedback questionnaires

A feedback questionnaire was sent out to Cloudy participants at the end of the study via the weekly study newsletter. The main focus of this questionnaire was to determine what factors participants liked and disliked about the study and the Cloudy app through a series of multiple choice questions and questions which allowed free-text responses. The results of these questionnaires were analysed using RStudio to complete the statistical analysis and Microsoft Excel to analyse the free-text responses. The themes included in the feedback questionnaire were: ease of use of the app, amount of time taken to complete data entry on a daily basis and throughout the study, reasons for taking part, overall opinion of the study, and opinions on the Cloudy app. Participants were divided into groups relating to the amount of time they remained in the study (obtained from information provided in the questionnaire). Beliefs about how worthwhile the study was, how easy the app was to use and how much time was required to take part were quantified for each group.

2.8 Statistical analysis

For each comparative analysis, data were tested for normality using the Shapiro-Wilk test. Samples that were too large for this test (>4000) or yielded a p-value of <0.05 were subjected to non-parametric tests. The Mann-Whitney U test was used to compare two groups (ages/weather beliefs of males and females in the total population). The mean number of days in study, days of data entry and percentage engagement of each cluster were compared as a group of 4 with the Kruskal-Wallis test, and in turn using the Mann-Whitney U test.

3. RESULTS

3.1 Descriptive analysis of the study population

A total of 13,256 people downloaded the Cloudy app. From this population, 1,267 individuals did not submit any symptom data over the entire period of data collection, bringing the total number of active participants to 11,989. In order to make sure that descriptive analysis could be carried out for each participant, it was necessary that they had submitted a baseline questionnaire. Participants who had not entered a baseline questionnaire (n=947) were therefore excluded. This left the total number at 11,042. The selection process is shown in Figure 2.

Figure 2: Flow-chart showing selection criteria

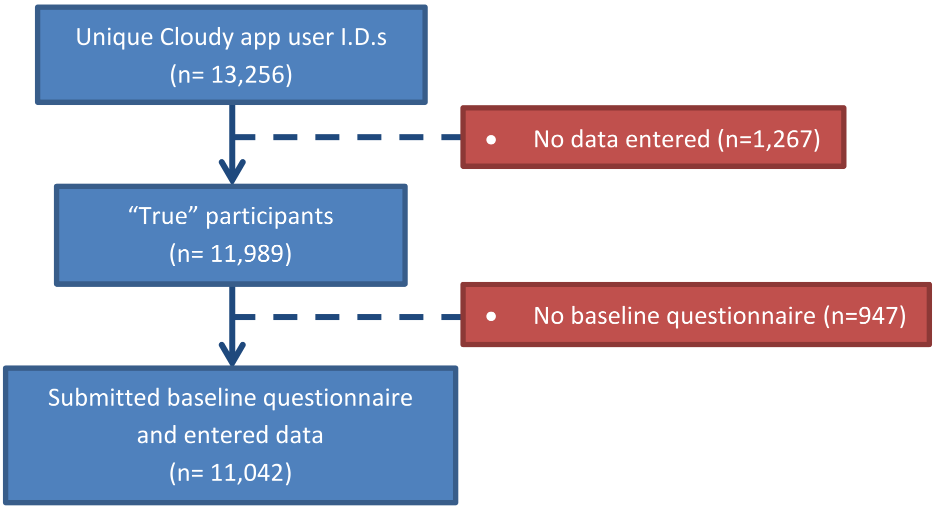


Figure 2: 13,256 individuals downloaded the Cloudy app and received unique user identification (user I.D.). Individuals who did not enter any data were excluded. Only those who submitted a baseline questionnaire were analysed (n=11,042).

The average age of the study participants was 49 and the majority of participants were female (80.35%). 11 participants did not provide a year of birth and 336 provided a year of birth that was impossible. Participants (n=28) who provided what appeared to be a date of birth in day/month/year form had the last two digits converted into a year (e.g. 4667=1967). Participants who provided what appeared to be the last two digits of a year (n=46) had this converted into a year (e.g. 98=1998). Table 2 outlines the general characteristics of the study population. The Shapiro-Wilk test for normality confirmed that the ages were not normally distributed. Males were found to be significantly older than females ($p<0.0001$) by the Mann-Whitney U test.

Table 2: Characteristics of the study population

	Total Population	**Female**	**Male**
***n* (%)**	11042	8872 (80.35)	2170 (19.65)
Mean age (±S.D.)	48.98 (±13.31)	47.22 (±13.29)	51.74 (±13.26)
Median age (I.Q.R.)	49 (40-59)	47 (38-57)	52 (43-62)

(S.D. = standard deviation; I.Q.R. = interquartile range)

There were 9 options for diagnosis in the baseline questionnaire. The free-text answers to "other" diagnoses were analysed and as there were a significant number of participants diagnosed with hypermobility syndromes (including Ehlers-Danlos syndrome) this was added as a new diagnostic group. The most common diagnosis was arthritis (unspecified type) with 34.8% participants selecting this option. This was followed by chronic widespread pain/fibromyalgia (26.47%) and the specified types of arthritis – osteoarthritis (22.52%) and rheumatoid arthritis (18.51%). Participants were able to select more than one diagnosis; hence the total number of diagnoses is greater than 11,042. The number of participants with each diagnosis is shown in Figure 3.

Figure 3: Bar chart showing the number of participants with each diagnosis

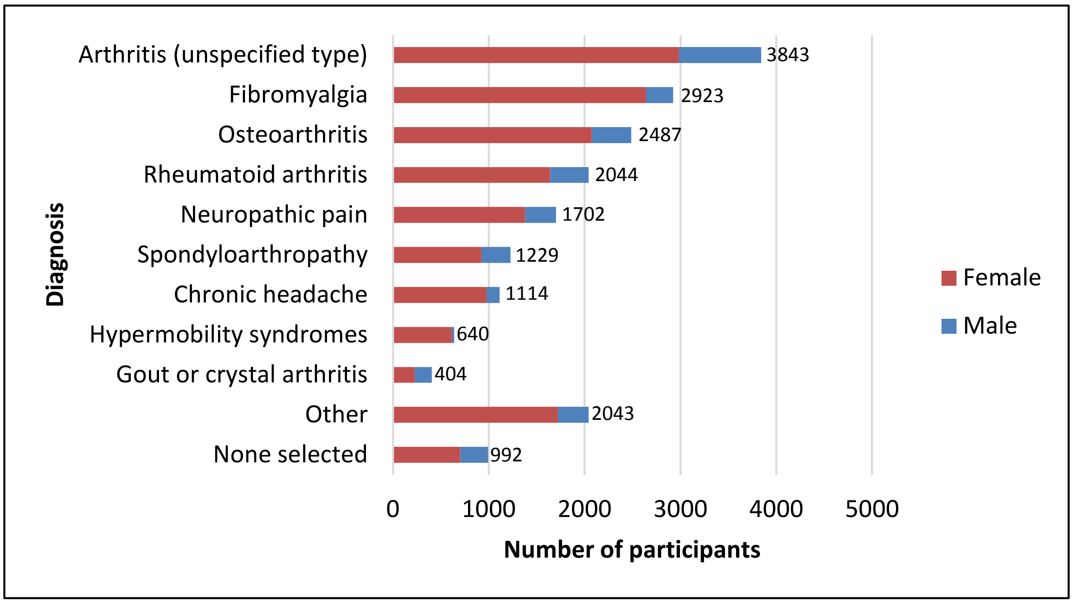


Figure 3: Bars represent the total number of participants with each diagnosis. The proportion of females is shown in red and males in blue.

The most common diagnosis in the female population was fibromyalgia/chronic widespread pain (39.95%) followed by unspecified arthritis (33.62%). In males, the most common diagnosis was unspecified arthritis (39.63%). Proportionally, more females were diagnosed with each condition – except for spondyloarthropathies and gout, where a higher relative proportion of the males in the study had these diagnoses.

Next, the sites at which participants experienced pain were analysed. 75 participants did not select any of the given pain sites, nor the option for multiple pain sites or all over pain. When those who had reported experiencing pain all over were removed (n=2,052), participants who had selected the given option for pain at multiple sites was found to be 3,220 (F=2675, M=545). In addition, there were 4,676 (F=3682, M=994) users who did not select this option but selected more than one given pain site – which gave a total of 7,896 participants experiencing pain at multiple sites but not all over. 1,033 participants (9.36%) had pain at a single site. This is summarised in Table 3.

Table 3: Pain sites

Pain site	Total, *n* (%)	Female, n (%)	Male, n (%)	Proportion of females %	Proportion of males %
Head	2130 (19.29)	1872 (87.89)	258 (12.1)	21.10	11.89
Face	808 (7.32)	727 (89.98)	81 (10.02)	8.19	3.73
Mouth/jaws	1722 (15.60)	1554 (90.24)	168 (9.76)	17.52	7.74
Neck/shoulders	6250 (56.60)	5221(83.54)	1029 (16.46)	58.85	47.42
Back	6489 (58.77)	5312 (81.86)	1177 (18.14)	59.87	53.24
Stomach/abdomen	1867 (16.91)	1613(86.40)	254 (13.60)	18.18	11.71
Hip	5660 (51.26)	4801 (84.82)	859 (15.18)	54.11	39.59
Knee	6912 (62.60)	5655 (81.81)	1257 (18.19)	63.74	57.93
Hand	6369 (57.68)	5335 (83.77)	1034 (16.23)	60.13	47.65
Feet	5203 (47.17)	4355 (83.70)	848 (16.30)	49.09	39.08
Multiple sites (but not all over)	7896 (71.51)	6357 (80.51)	1539 (19.49)	71.65	70.92
All over	2052 (18.58)	1804 (87.91)	248 (12.09)	20.33	11.42

Knee pain was the most common site of pain for females (63.74%), followed by hand pain (60.13%), back pain (59.87%) and neck/shoulder pain (58.85%). Similarly, the most common site of pain for males was knee pain (57.93%). This was followed by back pain (53.24%), hand pain (47.65%) and neck/shoulder pain (47.42%). A similar proportion of females and males experienced pain at more than one site (71.65% and 70.92% respectively). Proportionally, more females experienced pain all over compared with males (20.33% versus 11.42%).

3.2 Weather beliefs

Participants selected a number between 0 and 10 to represent the strength of their belief that weather affected pain. The modal score was 8 and the mean score was 7.22 (S.D. 2.11). Females thought that weather affected pain more than males, with the mean belief score of the females being significantly higher than that of the males ($p<0.0001$). 93.32% pf participants gave a rating of 5 or more and just 0.94% scored their belief as zero. The percentage of participants choosing each score is illustrated in Figure 4.

Figure 4: Graph showing participants' beliefs about how much the weather affects pain

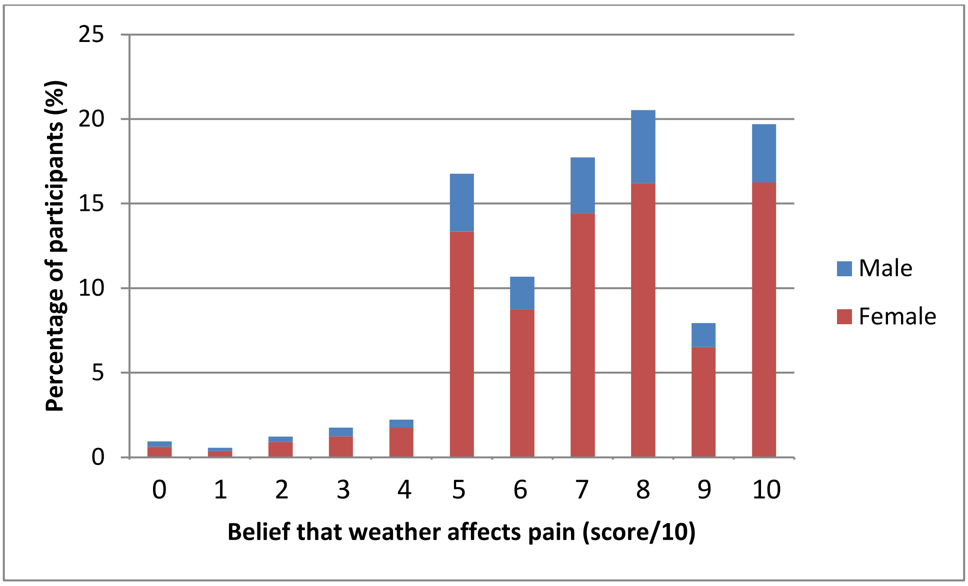


Figure 4: Bars represent the percentage of participants choosing each score to represent the strength of the association between weather and pain. Score 0 indicates no association and 10 indicates a strong association.

Participants could choose more than one option for the weather variables they believed to most affect their pain. The most common weather variable participants believed to affect their pain was damp/rain, with 72.73% participants choosing this option. Cold was the second most commonly selected parameter (65.48%). A similar percentage of participants believed that changes in barometric pressure and temperature affected their pain (34.62% and 33.96% respectively).

Answers to the free-text question (n=352) were analysed and broken down into other, unsure and no relationship. Two participants stated that all of the variables influenced their pain and therefore were included in the totals for each weather variable and excluded from the "Other" option. 27 participants had stated that they were unsure of (n=16) or did not believe in (n=11) any association between the weather and pain.

Only 0.1% of the Cloudy population believed that there was no relationship between the weather and their pain. Of those, there was almost an equal number of females and males (n=6 and n=5 respectively). However, this corresponds to a greater proportion of males from the overall male population. These findings are illustrated in Figure 5.

Figure 5: Graph showing the percentage of participants believing each weather variables affects pain severity

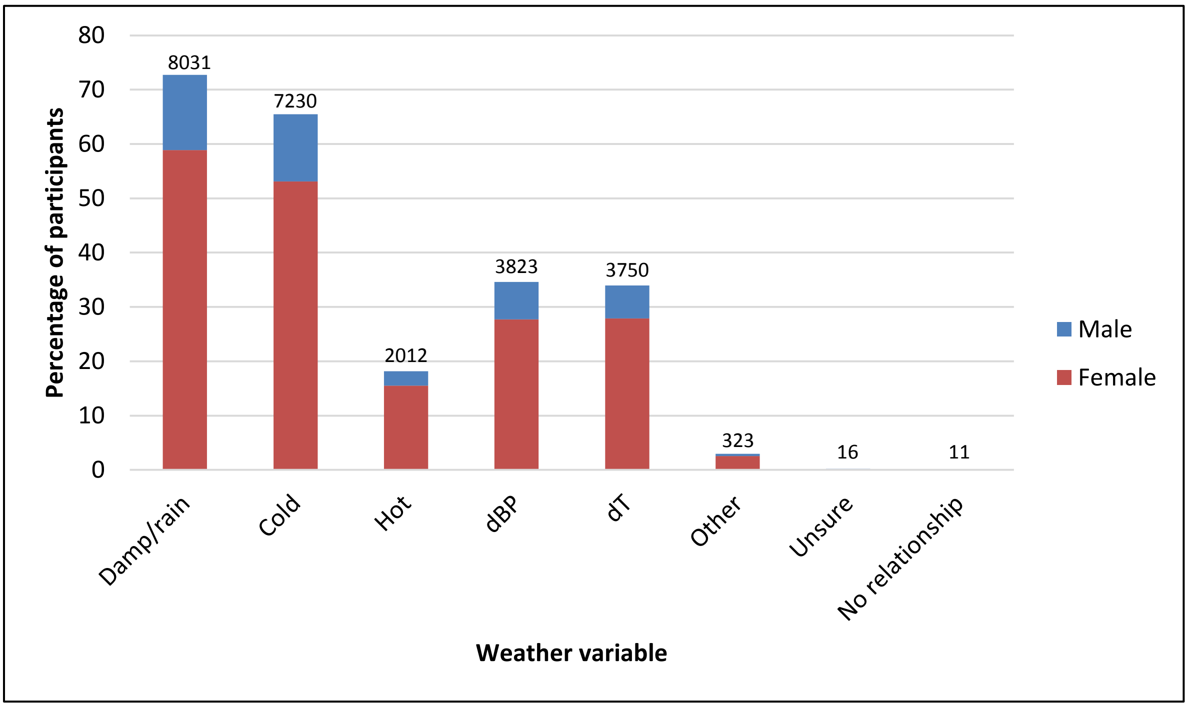


Figure 5: Bars represent the percentage of participants believing which weather variables have an effect on pain severity. Numbers above the bars are the total number of participants choosing that option. Those who were unsure/did not believe that a link between the weather and pain existed are also shown (dBP = change in barometric pressure; dT = change in temperature).

Within each diagnostic group (n=11), participants' beliefs about which variables most affected pain were investigated. The results are shown in Figure 6.

Figure 6: Graph showing the percentage of participants with beliefs about the contribution of each weather variable to their pain for each condition

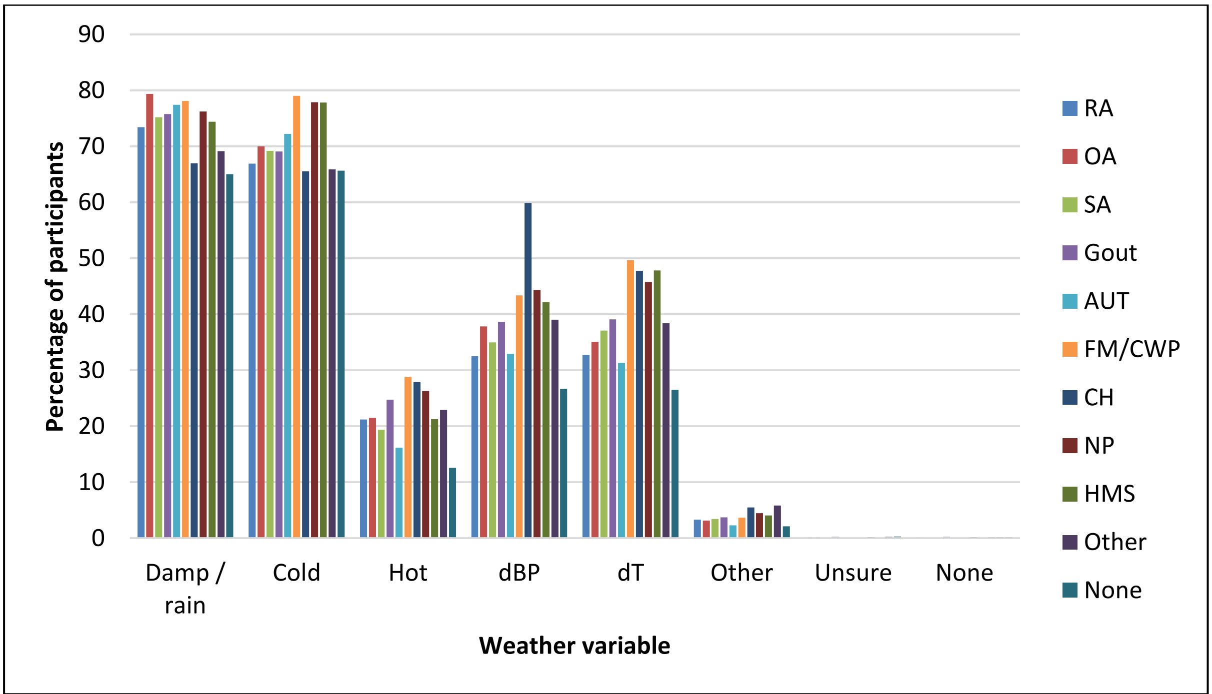


Figure 6: Bars represent the percentage of participants with each diagnosis believing each weather variable affected their pain (dBP = change in barometric pressure, dT = change in temperature, RA = rheumatoid arthritis, OA = osteoarthritis, SA = spondyloarthropathy, AUT = arthritis unspecified type, FM/CWP = fibromyalgia/chronic widespread pain, CH = chronic headache, NP = neuropathic pain, HMS = hypermobility syndromes).

Damp/rain was the most common weather variable believed to be associated with pain for most specified conditions (6/9) except fibromyalgia/chronic widespread pain, hypermobility syndromes and neuropathic pain – where slightly more participants thought cold had more of an effect. For those with specified diagnoses, damp/rain affected people suffering with osteoarthritis the most (79.33%) and those with chronic headache the least. Of the given weather variables, heat was thought to be least likely to affect pain. Changes in barometric pressure were believed to affect participants with chronic headache (59.87%) more than any other condition, and had least effect on those with rheumatoid arthritis and unspecified arthritis.

Changes in temperature were thought to influence pain in mainly in those with fibromyalgia (49.67%), chronic headache (47.76%) and neuropathic pain (45.77%).

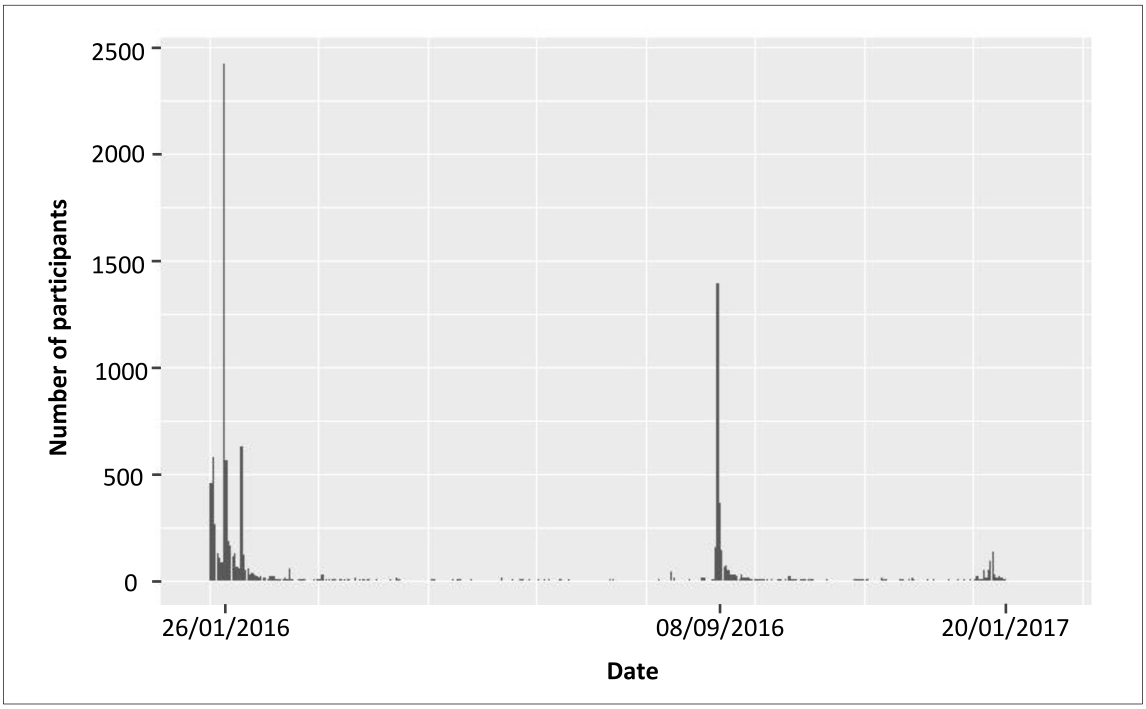
3.3 Enrolment analysis

The entire study period (20/01/16 – 20/04/17) lasted 457 days including the first and last possible days of data entry. Participants could enrol during the recruitment period of 20/01/16 – 20/01/17 (367 days). The day with the largest number of people enrolling (n=2423) was 26th January 2016, followed by 8th September 2017 (n=1395). These two dates coincided with press conferences and TV appearances to publicise the Cloudy study.

There were 16 days during the recruitment period on which no one enrolled in the study. The mean number of participants enrolling on each day in the study was 24 (S.D. 142.95) and the median number of enrolments per day was 4 (I.Q.R. 1-8). Figure 7 shows the pattern of participant enrolment over the enrolment period.

Figure 7: Graphs showing a) the number of participants enrolling on each day of the Cloudy study period and b) cumulative enrolment

a)



b)

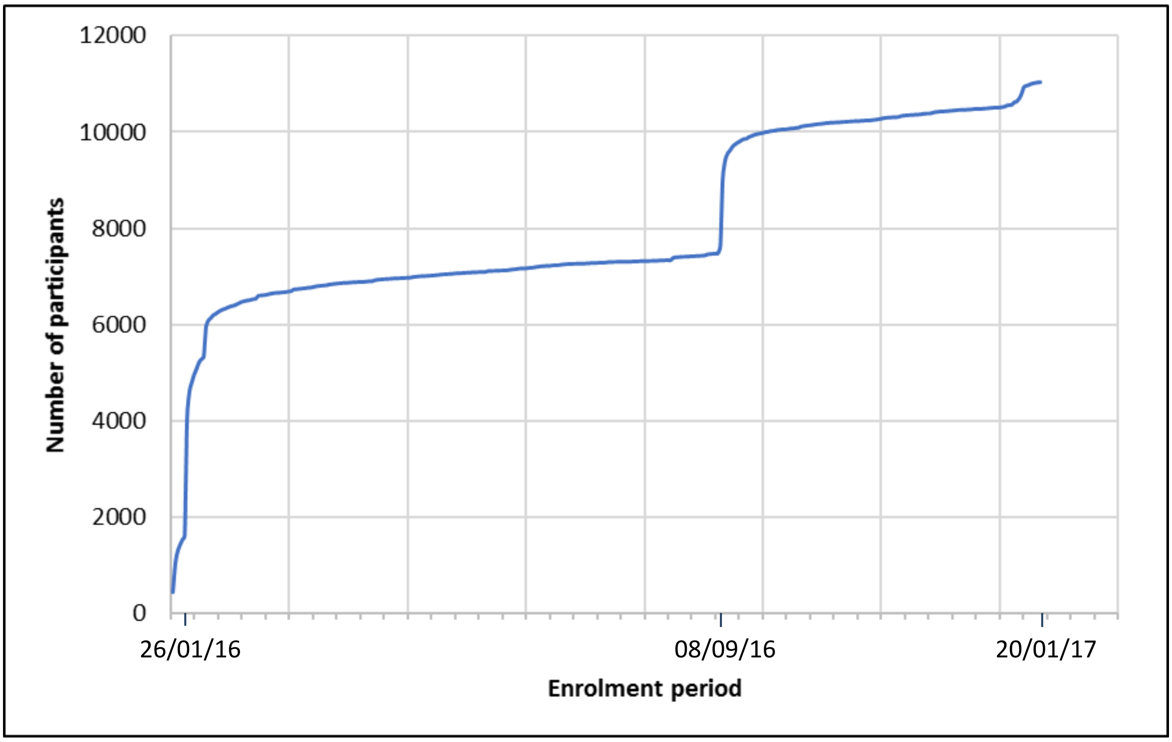


Figure 7: Graphs showing participant enrolment: 7a) bar-graph showing the number of participants enrolling on each day of the enrolment period (367 days); 7b) cumulative frequency curve showing the number of participants enrolling throughout the enrolment period.

Most participants stated that they found out about the Cloudy study through TV (50.78%). Social media was also important for publicising the study to a broader audience, with 19.44% of participants finding out about Cloudy through Facebook, Twitter, LinkedIn or other social media sites. This is illustrated in Figure 8.

Figure 8: Graph showing the means by which participants found out about the Cloudy study

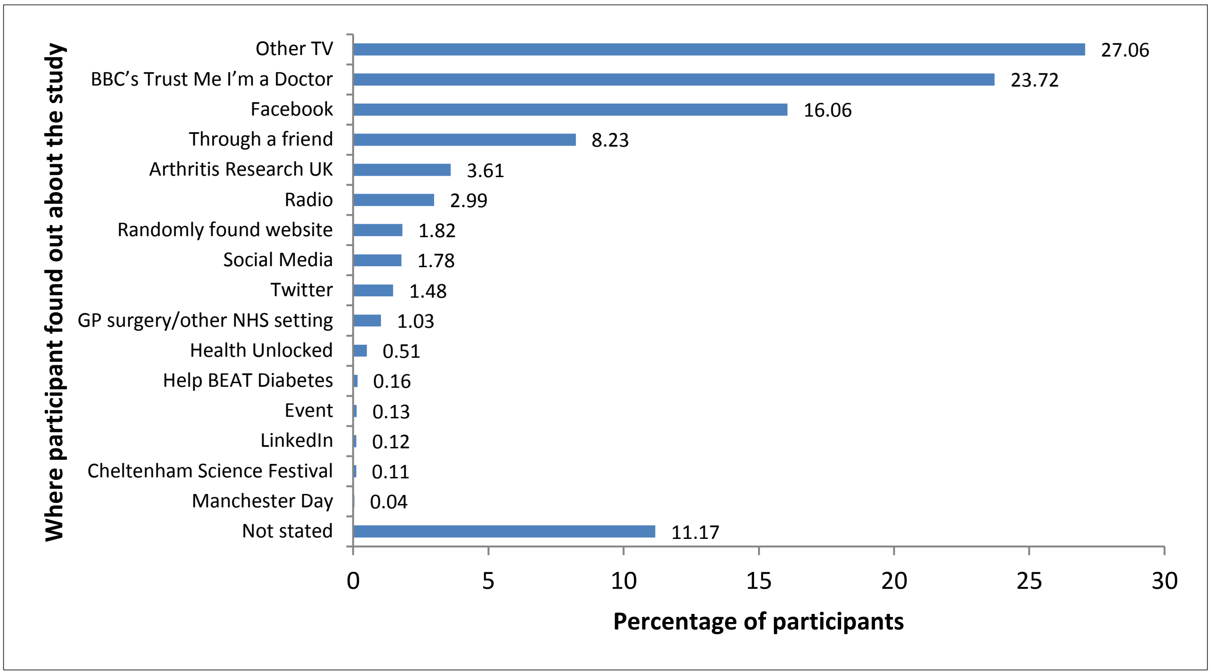


Figure 8: Bars represent the percentage of participants finding out about the study through per mode of advertisement.

3.4 Engagement patterns

456 (4.13%) users enrolled on the first day of the study and theoretically could have entered data for all 457 days. However, as data entry was only possible until 12 noon on the final day, it was more accurate to reduce this total number by 1 day to 456 days. The highest number of days of data entry was 453 (n=1) which equates to 99.34% engagement; whilst the lowest number of days of data entry was 1 (n= 2109, 19.10%).

The general engagement of the study population as a whole is summarised in Table 4. The mean number of days that participants stayed in the study (from enrolment to date of last data entry) was 72 and the mean number of days that participants entered data was 42. The mean overall percentage engagement was 13.01% – determined by the number of days a user entered data for over the entire period of time they theoretically could have entered data (see section 2.6).

Table 4: Data entry and engagement of the study population

Days of possible data entry		Days in study		Days of actual data entry		% engagement	
Mean (±S.D.)	**Median (I.Q.R.)**	**Mean (±S.D.)**	**Median (I.Q.R.)**	**Mean (±S.D.)**	**Median (I.Q.R.)**	**Mean (±S.D.)**	**Median (I.Q.R.)**
357.71 (±120.33)	442 (224-450)	72.23 (±109.61)	18 (3-96)	42.35 (±77.57)	8 (2-42)	13.01 (±22.00)	2.49 (0.66-13.84)

(S.D. = standard deviation; I.Q.R. = interquartile range)

3.5 Clusters of engagement

The Hidden Markov Model was applied to the entire dataset using methodology adapted from previous preliminary analyses of Cloudy engagement[34,35]. The population was divided into 4 clusters: 1, 2, 3 and 4. There were 753 participants who were not assigned to a cluster and this error will be explored by the study team in future analyses. Hence the clustering analysis includes 10,289 participants.

Cluster 3 had the highest percentage engagement (41.91%), with users in this group tending to stay in the study longer (mean 199.51 days) and enter data on an average of 139.81 days (mean). Thus, cluster 3 was regarded as the "high" engagement cluster. The Mann-Whitney U test demonstrated that participants in the high engagement cluster had significantly more days of data entry and higher percentage engagement than the other clusters ($p<0.0001$). Cluster 4 was the "low" engagement cluster with an average (mean) of 26.92 days in the study, 4.79 days of data entry and 1.55% engagement. Cluster 1 represented those who were between the low and high engagement clusters and were termed "moderate" engagers. The users who demonstrated the lowest levels of engagement, i.e. downloaded the app and entered data only once or a few times, were categorized into the "tourist" group. Tourists had the lowest percentage engagement at 0.40%, stayed in the study for an average of 3.44 days and entered data on 1.27 days. This is summarised in Table 5.

Table 5: Engagement of each cluster

Cluster (number)	Mean days in study (±S.D.)	Mean days of data entry (±S.D.)	Mean % engagement (±S.D.)
High (3)	199.51 (±126.62)	139.81 (±101.15)	41.91 (±26.38)
Moderate (1)	44.34 (±60.86)	15.25 (±13.61)	5.44 (±5.98)
Low (4)	26.92 (±65.24)	4.79 (±14.58)	1.55 (±4.03)
Tourist (2)	3.44 (±8.09)	1.27 (±2.71)	0.40 (±0.79)

(S.D. = standard deviation; I.Q.R. = interquartile range)

The characteristics of each cluster were analysed (Table 6) and it was found that those in the high engagement cluster were predominantly female (82.48%) and the ratio of females to males in this group was greater than that of all other clusters. The Shapiro-Wilk test for normality yielded p-values of <0.05 for the ages in all clusters – indicating that the data were non-normal. Three Mann-Whitney U tests and a Kruskal-Wallis test showed that participants in the high engagement cluster were significantly older than those in the other 3 clusters (p<0.0001) with a median age of 57. The cluster with the most participants was the moderate engagement cluster with 35.56% of users falling into this category (Figure 9).

Table 6: Characteristics of each engagement cluster

Cluster	***n* (%)**	**Female, *n* (%)**	**Male, *n* (%)**	**Mean age (±S.D.)**	**Median age (I.Q.R.)**	**Mean belief score (±S.D.)**	**Median belief score (I.Q.R.)**
High (3)	2802 (25.38)	2311 (82.48)	491 (17.52)	52.70 (±12.50)	54 (44-62)	7.24 (±2.12)	7 (6-9)
Moderate (1)	3927 (35.56)	3227 (82.17)	700 (17.83)	47.11 (±13.00)	47 (38-56)	7.29 (±2.05)	7 (6-9)
Low (4)	1817 (16.46)	1440 (79.25)	377 (20.75)	47.01 (±13.35)	47 (38-57)	7.25 (±2.11)	7 (6-9)
Tourist (2)	1743 (15.79)	1346 (77.22)	397 (22.78)	47.32 (±13.92)	47 (37-57)	6.27 (±2.20)	7 (5-8)

(S.D. = standard deviation; I.Q.R. = interquartile range)

The average score that participants in each cluster gave to what extent they believed there was a relationship between the weather and pain (0-10) was calculated and the Mann-Whitney U test was used to compare each group individually and the Kruskal-Wallis test was used to compare all 4 groups. There was no significant difference between the belief scores of the high, moderate and low engagement clusters (p<0.0001). However, the tourist cluster gave significantly lower scores than the high, moderate and low engagement clusters, with an average score of 6.27 (p<0.0001).

Figure 9: Graph showing the number of participants in each cluster

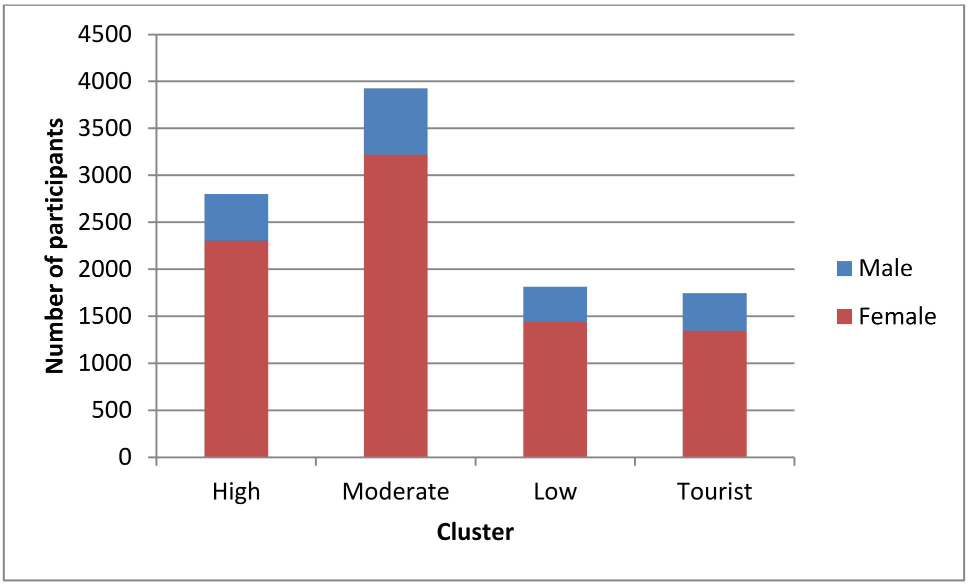


Figure 9: Bars show the number of participants in each cluster. Those who were not designated a cluster are not shown in the graph (n=753).

3.6 6-month study participation

Those who enrolled after 20/10/16 (n=861) would not have been able to participate for 6 full months as the last day of data entry fell before this, therefore they were not analysed. Of those who enrolled on or before this date (n=10,181), 1,605 stayed engaged for 6 months or more (calculated by adding 6 months to the day they enrolled and subtracting this date from the last day they entered data). 19 participants stayed in the study for exactly 6 months and 1,586 participants stayed engaged for over 6 months. Therefore, 15.76% stayed engaged for at least 6 months. 8,584 participants did not achieve the full 6 months, but 85 of these did manage to stay in the study until they were within 7 days of their 6-month mark. This is summarised in Table 7.

Table 7: Length of time participants stayed in the study with regards to the 6-month target

	Number	Days in study		Days of actual data entry		% engagement	
	n (%)	Mean (±S.D.)	Median (I.Q.R.)	Mean (±S.D.)	Median (I.Q.R.)	Mean (±S.D.)	Median (I.Q.R.)
6 months	19 (0.19)	182.95 (±0.23)	183 (183-183)	159.53 (±20.59)	166 (153-171)	40.81 (±10.16)	37.58 (35.78-39.78)
More than 6 months	1586 (15.58)	301.17 (±99.57)	256 (218-418.3)	183.85 (±115)	167 (99.75-253.3)	49.33 (±28.05)	45.56 (28.22-74.46)
Less than 6 months	8576 (84.24)	32.25 (±44.95)	10 (2-45)	17.37 (±28.32)	5 (2-19)	5.11 (±8.56)	1.36 (0.45-5.56)

(S.D. = standard deviation; I.Q.R. = interquartile range)

Those who stayed in the study for more than 6 months had a significantly higher percentage engagement than those who stayed in the study for less than 6 months ($p<0.0001$). There was no significant difference in the number of days of data entry or percentage contribution between the group who stayed in for exactly 6 months and the group that stayed in for more than 6 months ($p>0.05$).

3.7 Feedback questionnaires

529 participants returned the feedback questionnaire, which equates to 4.45% of the total participating study population with or without a submitted baseline questionnaire (n=11,989). Similar to the overall study population, there was a significantly greater proportion of females in this sample than males. The general characteristics of the sample returning the feedback questionnaire are summarised in Table 8.

Table 8: Characteristics of the population who returned a feedback questionnaire

Participants	Female, *n* (%)	Male, *n* (%)	Mean age (±S.D.)	Median age (I.Q.R.)
529	463 (87.52)	66 (12.48)	55.27 (±12.70)	56 (47-65)

(S.D. = standard deviation; I.Q.R. = interquartile range)

Those returning the feedback questionnaire were separated into groups (n=6) depending on how long they remained in the study. These groups were determined by answers to a multiple-choice question in the questionnaire with options: more than one day but less than one week; one week but less than one month; 1-2 months; 3-5 months; 6-12 months; and 12 months or more. 96.40% of responders to the feedback questionnaire stayed in the study for at least 3 months, with 77.50% of responders staying engaged for over 6 months. 8 of the responders to the feedback questionnaire had dropped out before 1 month of study participation. This is summarised in Table 9.

Table 9: Length of time participants who returned a feedback questionnaire stayed in the study

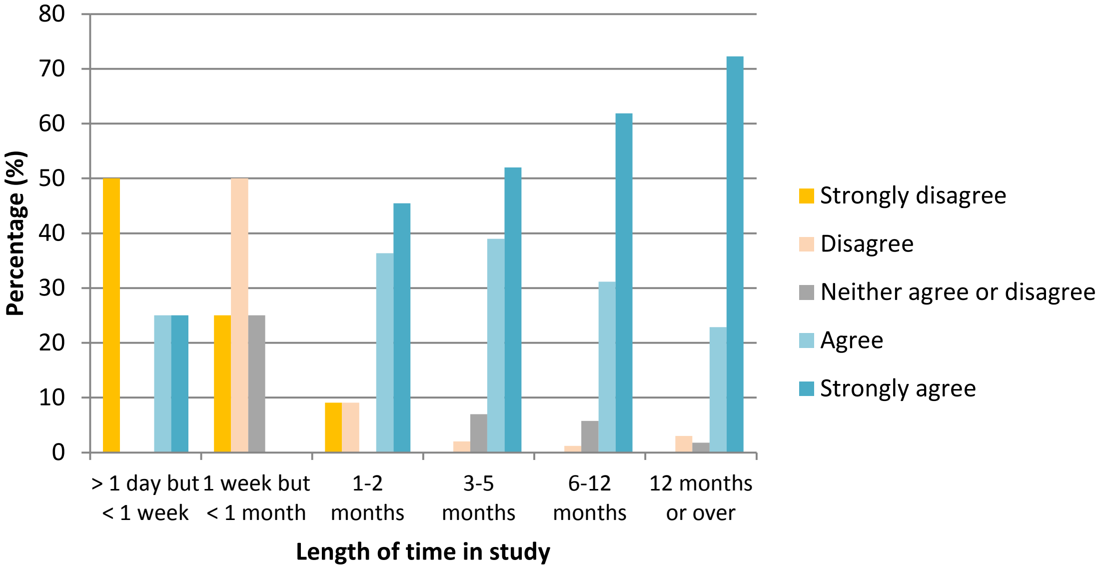
Length of time	>1 day but < 1 week	1 week but < 1 month	1-2 months	3-5 months	6-12 months	12 months or over
***n* (%)**	4 (0.76)	4 (0.76)	11 (0.19)	100 (18.90)	244 (46.12)	166 (3.38)

Overall, feedback was very positive, with 96.59% responding that they felt it was worthwhile to take part in the Cloudy study and 95.84% indicating that they would recommend a similar study to a suitable friend. The most common reason for taking part in the study was to help find a relationship between the weather and pain (79.58%). This was followed by wanting to contribute to research that might benefit participants directly (56.71%) and research that might benefit others (49.15%). 28.92% of responders participated in the study so that they could track their symptoms through the app and 14.56% were motivated to take part in the study because they had an interest in technology.

The longer participants had stayed in the study, the more likely they were to strongly agree that using the Cloudy app was easy and quick. This is illustrated in Figure 10.

Figure 10: Participant beliefs that it was a) easy and b) quick to input symptom data into the app

a)



b)

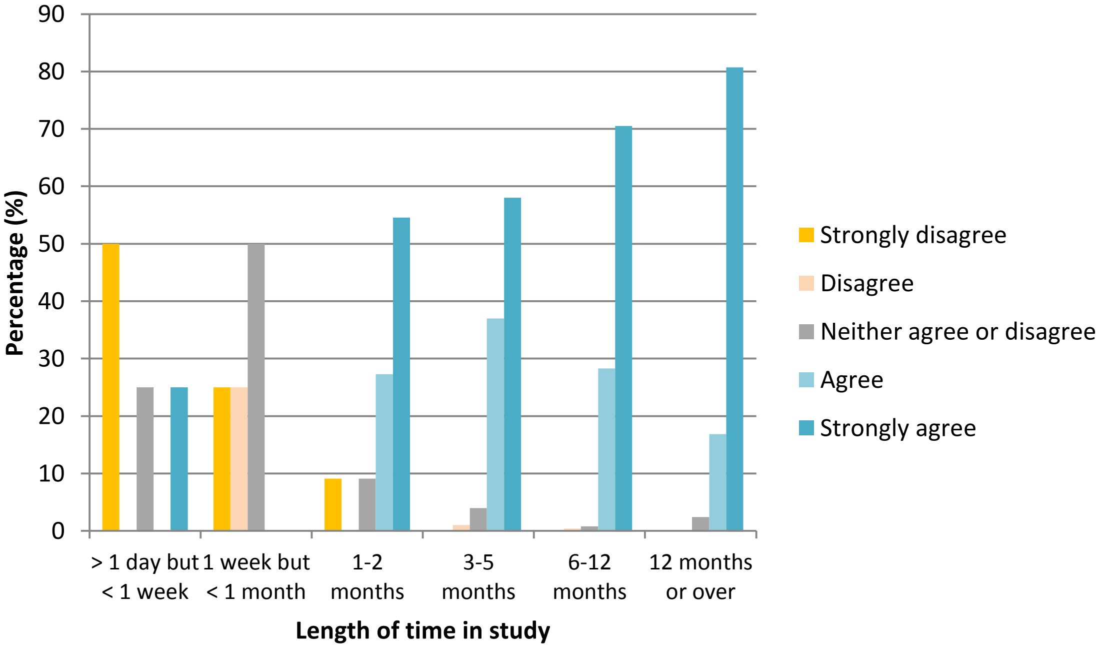


Figure 10: Bar graphs showing the percentage of responders to the feedback questionnaire who thought it was easy (graph a) and quick (graph b) to use the Cloudy app. The number of participants in each group can be found in Table 9.

The main themes from the free-text answers to the feedback questionnaire are shown in Table 10. Whilst many found the app easy and quick to use, there were others who thought the app was too complex and time consuming.

Table 10: Feedback from free-text comments of the feedback questionnaire

Positive feedback		Negative feedback	
App	**Study**	**App**	**Study**
• Simple to use • Quick to do • Attractive design • Daily reminder • Good technical help • Provided a way of monitoring symptoms • Diary feature	• Easy to enrol and enter data • Interesting • Newsletters useful • Responsive team • Well organised	• Too complex to use • Difficult to read and enter data on small screen • Problems with GPS and weather data collection • App dysfunction • Difficult to retrospectively enter symptom data	• Required too much time and commitment • Diet and stress level not addressed • Made users focus on pain more than usual • Choices not specific enough • Subjective measures were difficult to quantify
Reasons for taking part / disengaging		**Suggestions for improvement**	
• To contribute to research • To increase understanding of chronic pain and the weather • To benefit others suffering from chronic pain • Other commitments/life events • Issues with phone/app function		• More options for the exercise/activity level • Built-in step recorder • Facility to print out results • Bigger range of point-scales • Guide to using the app – videos, manuals • Facility to record data whilst abroad • Include more categories e.g. sleep, work, diet • Ability to enter different pain scores at multiple sites	

Those who dropped out of the study in the first week gave reasons such as being too busy, forgetting to fill in data and having technical difficulties with their phone or the app for stopping. Of the 4 participants who stayed in the study for 1 week but less than 1 month, 3 of them commented that they found using the app difficult or problematic, whilst 1 of them stated that they became too busy to continue using the app.

4. DISCUSSION

4.1 Principal findings

It was unsurprising that the majority of participants in the Cloudy study were females, as it is generally more likely for females to voluntarily take part in research studies than males[36,37]. Furthermore, the target population of the study were people suffering from chronic pain – a condition which predominantly affects people of female gender and of older age[3]. The average age of the study population was 49, yet chronic pain is usually more prevalent amongst those aged 65 and older[38]. Therefore, there may have been some sampling bias towards a younger population of chronic pain sufferers due to the nature of it being a smartphone study. As males in the study were significantly older than females, it would be worthwhile to find ways to appeal to younger males to obtain a wider age spread.

The most common participant diagnosis was unspecified arthritis, which could include a mixture of the multiple types of arthritides including osteoarthritis, rheumatoid arthritis, psoriatic arthritis etc. There was an issue in the initial study period where the option for osteoarthritis was not present in the baseline questionnaire, so it is possible that many of these individuals would have picked osteoarthritis had it been available at the time of questionnaire submission. This would likely take the total of participants suffering with osteoarthritis to more than that of the second most common diagnosis (fibromyalgia/chronic widespread pain). Both of these conditions have been investigated in the past with regards to the weather's effect on their severity[7,9,10,11] and it will be interesting to compare the final findings of the Cloudy study with the results of previous investigations as there will be a much larger sample sizes from which to draw more reliable conclusions. The most common pain site for both males and females was knee pain, which is most often the result of osteoarthritis[39]. As unspecified arthritis was the commonest diagnosis of the study population, a high prevalence of knee pain is unsurprising.

Damp and rain were the weather variables most believed to affect participants' pain according to the baseline questionnaire. In previous studies, damp has been perceived by participants as a principal factor in worsening pain severity[40]; but has actually appeared to have minimal effect according to weather data matched to severity scores[41]. Atmospheric pressure changes are often associated with increasing joint pain severity[41], yet participants in Cloudy believed damp/rain and cold to have more of an effect on their pain that pressure changes. As the sample size of Cloudy is much larger than that of previous investigations, and data has been collected over all four meteorological seasons, it will be interesting to explore this further when participants' actual pain scores as opposed to their initial beliefs are matched to weather data.

It is clear that using a smartphone app is an effective way of recruiting large number of subjects in a short space of time, with 4,059 of the population investigated in this report having enrolled in the first week. In addition to this, it has clearly demonstrated that smartphones allow continuous recruitment and enrolment of new subjects throughout a study period. Similar observations were made by the Asthma Mobile Health Study[33] where 43,949 downloads of their app were achieved in the first month. However, only 7,593 individuals actually participated in that study and although participant enrolment was high at the beginning of the study, it greatly decreased to between 300 and 400 downloads per month by the sixth month of the study period. An analogous phenomenon was observed in the Cloudy enrolment analysis, with fewer enrolments per month as the study progressed. It is possible that this may be a trend that will continue to manifest in large mHealth studies.

It is important to note that substantial numbers of participants enrolled on days of, or just after, a significant press release or large publicising event. This indicates that exposure through large media platforms is an efficacious way of recruiting large numbers of volunteers for mHealth studies. This had also been noted by the Asthma Mobile Health Study[33] where investigators attributed the initial surge of enrolment in their study to media publicity.

This study has demonstrated that a substantial number of participants stay engaged with an mHealth app and continuously enter data over time; with 15.76% of the sample used in the 6-month analysis (n=10,181) staying engaged for at least 6 months. This is an important finding because this equates to an ample sized study population in its own right (n=1,605) – which is encouraging for future mHealth studies which may require long-term monitoring and follow-up of participants. Furthermore, over one-quarter of those in this analysis were considered “high” engagers and over one-third were “moderate” engagers which is a promising result for future longitudinal mHealth studies.

A notable proportion of users (15.79%) were categorised as being “tourists” in that they only interacted with the app on one occasion or a handful of times over a short period of time. This is a relatively large number of users to be lost, but there has been some insight into why people drop out after such little time from responses to the feedback questionnaire. Some reasons given were forgetting, other commitments and technical difficulties. However, 77.50% of responders to the feedback questionnaire claimed to have stayed in the study for 6 months or more, which is a much greater proportion than did so for the sample population mentioned above (15.76%). Therefore, the feedback questionnaire was answered by a sample of the population who had engaged with the study for a longer period of time, with only 8 responses from users who dropped out before a full month of using the app. It would be useful to gather feedback from more participants who dropped out early to find out as much information as possible about which factors make individuals more likely to disengage.

4.2 Engagement enablers

Participant characteristics appear to be associated with engagement with mHealth studies. In this study, high engagement was associated with older age and female gender. Higher engagement in older age groups is something that has been noted previously[42] and may be a shared characteristic of engagement patterns in mHealth studies. The observation that females seemed to engage more than males may be due to this study population having a significant female majority, as other trials have observed higher mHealth engagement in males than females[43]. Previous investigations have found a relationship between employment status and mHealth study engagement, for example Min *et al*[23] found that unemployed women exhibited greater compliance. Employment status was not recorded in the Cloudy baseline questionnaire but it would be interesting to explore this element in future large-scale studies as the sample size of the aforementioned study[23] was small with just 30 participants.

Although there was an encouraging percentage of participants staying engaged for 6 months or more in Cloudy, a previous mHealth study[27] managed to achieve a much greater retention of participants over a 6-month period (93%). A potential reason for this could be that their app was directly beneficial to the study population in that it helped them to lose weight – whereas the Cloudy app was primarily for research. Hence, it is possible that actual and perceived benefits from using a mHealth app act as an incentive to continue engaging with it.

From the feedback questionnaire, it appeared that people who stayed in the study for longer periods of time believed that the app was easy to use and that entering data did not take up very much time. Although there were fewer low-engaging participants who returned the feedback questionnaire, these participants were more likely to think that the app was less easy to use or that it required too much time. This supports previous findings that mHealth apps must be simple to use and not be too time-consuming so that they remain attractive to participants[30,31].

Participants were provided with both personal and study population data visuals, a facility to track their own symptoms over time and a monthly study newsletter. These features were well received by participants according to the feedback questionnaires. In addition to features like these, it has been suggested that gamification of mHealth apps may be a strategy to retain participants over longer periods of time[44], where leader-boards and virtual rewards may act as incentives for participants to stay engaged.

4.3 Barriers to engagement

There are a number of things that contribute to participant disengagement. Firstly, age is a significant factor which must be addressed. It has been suggested that elderly populations may be less likely to engage with mHealth programmes because of limited or lack of access to relevant technologies, including smartphones[31]. There were less participants at the extremes of age enrolling in the Cloudy study than would have been predicted from the age spread of the UK chronic pain population according to the Health Survey for England[45]; and this is something that will need to be considered in further mHealth studies, with attempts made to appeal to both younger and older populations. Another barrier to using mobile apps is that a level of literacy is required which is not always possessed by members of the target population[31]. This was alluded to in some responses to the feedback questionnaire where participants said they found the app too complex to use or difficult to understand.

Interestingly, the observed inclination to lose participants early on is not only the case for mHealth research studies; rather it is something that affects health-related mobile apps in general. The Consumer Health Information Corporation (CHIC)[46] have reported that 26% of health apps are used only once by downloaders and that approximately 74% discontinue engaging with an app after ten uses.

Investigations focussing on user engagement and attrition in a pilot study for Cloudy[47] highlighted reasons for participants dropping out. A reduction in battery life and problems with phone storage caused by the app were important issues affecting participants who provided feedback in focus group meetings. These issues were rectified for the Cloudy study so were not as relevant for the attrition of the population analysed in this report – however they are key factors to bear in mind for future mHealth apps.

As participant engagement is generally measured by quantifying a user's data contribution, problems can arise when there are problems with data collection. This can mean that although participants may still be engaging with the study and or app, their data is not successfully recorded and so it may appear that they have become disengaged or dropped out. An exploratory study by Saeb *et al*[48] investigating factors influencing depressive symptoms via a mobile phone sensor highlighted reasons for data collection failure including:

issues with connection to the data server, participants not charging their phones and unavailable network connections over extended periods of time.

4.4 Strengths and weaknesses of the study

The study population has been successfully categorised into four groups depending on their engagement with the Cloudy app. The association with high engagement, female gender and older age has provided some insight into the participant characteristics which may affect mHealth study engagement. It would be interesting to investigate how employment status and mood affected engagement in this population like other mHealth studies have done[23,24] – although there is no recorded data for employment status, information about participants' mood is available.

However, as data was collected continuously throughout the study period, it is difficult to compare the engagement patterns of people who have been in the study for 12 months with people who have only been in it for a few weeks or months depending on when they enrolled and dropped out. For example, someone who engaged avidly at the beginning and then had a break for a while before returning to enter data would exhibit high average engagement for the initial period, but have a lower overall engagement.

Due to time limitations, it was not possible to investigate the isolated engagement patterns of participants' 6-month study participation periods. This is something that would be worthwhile to explore in future Cloudy analyses.

Missing data is something that will need to be addressed in future studies. As year of birth was entered as text, some participants gave impossible answers or no answer at all. In addition, their age had to be calculated by subtracting this number from 2017. Perhaps a drop-down menu of ages would be more appropriate – so users could select an accurate value and their exact age would be recorded as opposed to calculated.

The feedback questionnaire was advertised through the study newsletter and responded to by 4.45% of participants. This is a small number of replies, and considering that 77.50% of these participants had been active participants for over 6 months, they are not representative of the overall study population. Only 1.51% of responders to the feedback questionnaire had dropped out of the study within 1 month. This means that those participants who dropped out early on were not reliably represented in this sample. However, it does highlight that although these participants dropped out early, they must have still had some channel through which to access the questionnaire – either through the newsletter or perhaps from other participants.

4.5 Limitations of external validity

The individuals involved in this large cohort study represented a diverse group of participants suffering from chronic pain with regards to age, gender, site of pain, diagnosis, and familiarity with smartphone apps and research. However, these participants were similar in that almost all of them harboured some sort of belief that the weather was associated with pain and its severity, with only 0.94% of participants scoring the relationship between weather and pain as zero. Participants were encouraged to enrol in the study if they suffered from chronic pain regardless of their ideas regarding the weather and its effects; yet, as participants self-selected themselves into the study, it is likely that already holding beliefs that there was a link between the weather and pain would increase someone's likelihood of enrolling in the study. This is supported by findings that the mean score chosen to quantify the extent of the effect of the weather on pain was 7.22 and the modal score given was 8 – both scores corresponding to strong beliefs in an association.

It is evident from both the overall study population and the sample who returned the feedback questionnaire that there was some sampling bias towards females, especially those of older age. However, as mentioned previously, it is more likely for females rather than males to volunteer for health studies and this has been acknowledged in previous mHealth studies[48]. Likewise, chronic pain is something which is associated with female gender and older age[49]. Therefore, the sample investigated does represent a biased population with regards to the inclusion of participants sharing specific characteristics. In addition, the need to have a smartphone in order to enrol and continually engage in the study excluded certain individuals from taking part given that not everyone has the access to such technology.

5. CONCLUSION

It can be concluded from the successful enrolment of a sizeable study population in Cloudy with a Chance of Pain that smartphones offer a feasible way to collect large amounts of data over both short and long periods of time; and without the significant extra work and cost associated with traditional study recruitment processes. Furthermore, it is evident that people do stay engaged with mHealth applications over extended periods, with 1,605 of the 11,042 participants studied actively participating in the Cloudy study for 6 months or more. However, although this may be true overall, different patterns of participant engagement exist within mHealth studies. General and specific factors contribute to this, such as participants' perception and interest in the study, accessibility and convenience of the app, support offered by the study team and personal attributes for example gender, age, educational level and memory.

By understanding the way in which people engage with mHealth studies, future large-scale smartphone-based studies can be designed to encourage maximum participant recruitment and sustained engagement by adapting apps to suit diverse sample populations. This can be achieved by including features such as more robust reminders and/or alarms, large buttons and text, simple and attractive graphics, and user guides – whilst trying to limit the amount of time required to participate on a daily basis and the extent of battery usage and phone memory.

6. REFERENCES

1) Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *Eur J Pain* 2006;10(4): 287–333.

2) Merskey H, Bogduk N. Classification of chronic pain: descriptions of chronic pain syndromes and definitions of pain terms. Seattle, WA: IASP Press, 1994.

3) van Hecke O, Torrance N, Smith BH. Chronic pain epidemiology and its clinical relevance. *Br J Anaesth* 2013;111(1): 13-18.

4) Redelmeier, DA, Tversky A. On the belief that arthritis pain is related to the weather. *Proc Natl Acad Sci* 1996;93(7): 2895-2896.

5) Smedslund G, Hagen KB. Does rain really cause pain? A systematic review of the associations between weather factors and severity of pain in people with rheumatoid arthritis. *Eur J Pain* 2011;15(1): 5–10.

6) Drane D, Berry G, Bieri D, McFarlane AC, Brooks P. The Association Between External Weather Conditions and Pain and Stiffness in Women With Rheumatoid Arthritis. *J Rheumatol* 1997;24(7): 1309-1316.

7) Timmermans EJ, van der Pas S, Schaap LA, Sánchez-Martínez M, Zambon S, Peter R, Pedersen NL, Dennison EM, Denkinger M, Castell MV, Siviero P, Herbolsheimer F, Edwards MH, Otero A, Deeg DJ. Self-perceived weather sensitivity and joint pain in older people with osteoarthritis in six European countries: results from the European Project on OSteoArthritis (EPOSA). *BMC Musculoskelet Disord* 2014;15:66.

8) Wilder FV, Hall BJ, Barrett JP. Osteoarthritis pain and weather. *Rheumatology (Oxford)* 2003;42: 955–958.

9) McAlindon T, Formica M, Schmid CH, Fletcher J. Changes in barometric pressure and ambient temperature influence osteoarthritis pain. *Am J Med* 2007;120(5): 429–34.

10) Smedslund G, Eide H, Kristjansdottir ÓB, Nes AA, Sexton H, Fors EA. Do weather changes influence pain levels in women with fibromyalgia, and can psychosocial variables moderate these influences? *Int J Biometeorol* 2014;58(7): 1451–1457.

11) Hagglund KJ, Deuser WE, Buckelew SP, Hewett J, Kay DR. Weather, beliefs about weather, and disease severity among patients with fibromyalgia. *Arthritis Care Res* 1994;7(3): 130–135.

12) Beilken K, Hancock MJ, Maher CG, Li Q, Steffens D. Acute Low Back Pain? Do Not Blame the Weather — A Case-Crossover Study. *Pain Med* 2017;18(6): 1139-1144.

13) Edefonti V, Bravi F, Cioffi I, Capuozzo R, Ammendola L, Abate G, Decarli A, Ferraroni M, Farella M, Michelotti A. Chronic pain and weather conditions in patients suffering from temporomandibular disorders: a pilot study. *Community Dent. Oral Epidemiol* 2012;40: 56–64.

14) Smedslund G, Mowinckel P, Heiberg T, Kvien TK, Hagen KB. Does the weather really matter? A cohort study of influences of weather and solar conditions on daily variations of joint pain in patients with rheumatoid arthritis. *Arthritis Rheum* 2009;61: 1243–1247.

15) GPS. *National Geographic Society* 2012. Available from: https://www.nationalgeographic.org/encyclopedia/gps/ [Accessed 24th May 2017].

16) WHO Global Observatory for eHealth. *New horizons for health through mobile technologies* Geneva: World Health Organization; 2011: 112.

17) Sama PR, Eapen ZJ, Weinfurt KP, Shah BR, Schulman KA. An evaluation of mobile health application tools. *JMIR Mhealth Uhealth* 2014;2(2): e19.

18) van Velthoven MH, Car J, Zhang Y, Marušić A. mHealth series: new ideas for mHealth data collection implementation in low- and middle-income countries. *J Glob Health* 2013;3:020101.

19) Number of smartphone users worldwide from 2014 to 2020 (in billions). Available from: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [Accessed 24th May 2017].

20) Hayden EC. Mobile-phone health apps deliver data bounty: smartphone programs allow researchers to recruit large numbers of participants and monitor them in real time. *Nature* 2016;531(7595): 422–423.

21) Bitton A, Poku M, Bates DW. Policy context and considerations for patient engagement with health information technology. In: Grando MA, Rozenblum R, Bates D (eds) Information technology for patient empowerment in healthcare. Berlin: Walter de Gruyter GmbH & Co KG, 2015.

22) Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, Doerr M, Pratap A, Wilbanks J, Dorsey ER, Friend SH, Trister AD. The mPower study, Parkinson disease mobile data collected using ResearchKit. *Sci Data* 2016;3: 160011.

23) Min YH, Lee JW, Shin YW, Jo MW, Sohn G, Lee JH, Lee G, Jung KH, Sung J, Ko BS, Yu JH, Kim HJ, Son BH, Ahn SH. Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. *J Med Internet Res* 2014;16(5): e135.

24) Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. *BMJ Innov* 2016;2: 14-21.

25) Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: Review of the emerging field and proof of concept study. *J Ment Health* 2011;20(6): 509-524.

26) Kirwan M, Vandelanotte C, Fenning A, Duncan MJ. Diabetes Self-Management Smartphone Application for Adults With Type 1 Diabetes: Randomized Controlled Trial. Eysenbach G, ed. *J Med Internet Res* 2013;15(11): e235.

27) Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a Smartphone Application for Weight Loss Compared to Website and Paper Diary: Pilot Randomized Controlled Trial. Eysenbach G, ed. *J Med Internet Res* 2013;15(4): e32.

28) Quintiliani LM, Mann DM, Puputti M, Quinn E, Bowen DJ. Pilot and Feasibility Test of a Mobile Health-Supported Behavioral Counseling Intervention for Weight Management Among Breast Cancer Survivors. *JMIR Cancer* 2016;2(1): e4.

29) Painter S, Ditsch G, Ahmed R, Hanson NB, Kachin K, Berger J. Retrofit Weight-Loss Outcomes at 6, 12, and 24 Months and Characteristics of 12-Month High Performers: A Retrospective Analysis. Eysenbach G, ed. *JMIR mHealth and uHealth*. 2016;4(3): e101.

30) Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1): e11.

31) Martin, T. "Assessing mHealth: Opportunities and Barriers to Patient Engagement." *J Health Care Poor Underserved* 2012;23(3): 935-941.

32) Anhøj J, Møldrup C. Feasibility of Collecting Diary Data From Asthma Patients Through Mobile Phones and SMS (Short Message Service): Response Rate Analysis and Focus Group Evaluation From a Pilot Study. *JMIR* 2004;6(4): e42.

33) Chan YY, Wang P, Rogers L, Tignor N, Zweig M, Hershman SG, Genes N, Scott ER, Krock E, Badgeley M, Edgar R, Violante S, Wright R, Powell CA, Dudley JT, Schadt EE. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. *Nat Biotechnol* 2017:35(4): 354-362.

34) Selby D. Cloudy with a chance of engagement. *Significance* 2017;14:10-11.

35) Druce KL, McBeth J, van der Veer SN, Selby DA, Vidgen B, Georgatzis K, Lakshminarayana R, Chowdhury A, Schultz DM, Sanders S, Sergeant JC, Dixon WG. Recruitment and ongoing engagement in a UK smartphone study examining the association between weather and pain. Submitted 2017.

36) Lobato L, Bethony J M, Pereira FB, Grahek SL, Diemert D, Gazzinelli MF. Impact of gender on the decision to participate in a clinical trial: A cross-sectional study. *BMC Public Health* [Internet] 2014;14(1): 1156. doi:10.1186/1471-2458-14-1156

37) Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of Smartphone Applications and Wearable Devices for Tracking Physical Activity Data. *JAMA* 2015;313(6): 625-626.

38) Reid MC, Eccleston C, Pillemer K. Management of chronic pain in older adults. *BMJ* 2015;350: h532.

39) Nguyen U-SDT, Zhang Y, Zhu Y, Niu J, Zhang B, Aliabadi P, Felson, DT. Increasing Prevalence of Knee Pain and Symptomatic Knee Osteoarthritis. *Ann Intern Med* 2011;155(11): 725-732.

40) Jamison RN, Anderson KO, Slater MA. Weather changes and pain: perceived influence of local climate on pain complaint in chronic pain patients. *Pain* 1995;61(2): 309–315.

41) Brennan SA, Harney T, Queally JM, O'Connor McGoona J, Gormley IC, Shannon FJ. Influence of weather variables on pain severity in end-stage osteoarthritis. *Int Orthop* 2012;36: 643–646.

42) Goyal S, Morita PP, Picton P, Seto E, Zbib A, Cafazzo JA. Uptake of a consumer-focused mHealth application for the assessment and prevention of heart disease: the <30 days study. *JMIR Mhealth Uhealth* 2016;4(1): e32.

43) Zhang X, Guo X, Lai K, Guo F, Li C. Understanding gender differences in m-health adoption: a modified theory of reasoned action model. *Telemed J E Health* 2014 Jan;20(1): 39–46.

44) Miller AS, Cafazzo JA, Seto E. A game plan: Gamification design principles in mHealth applications for chronic disease management. *Health Informatics J* 2016;22(2): 184–193.

45) Bridges S. Health Survey for England 2011: Vol. 1 Chapter 9: Chronic Pain. 2012

46) Consumer Health Information Corporation. Motivating Patients to Use Smartphone Apps [Internet]. USA: Consumer Health Information Corporation; 2011 Apr [cited 2017 May 26]. Available from: http://www.consumer-health.com/motivating-patients-to-use-smartphone-health-apps/

47) Reade S, Spencer K, Sergeant J, Sperrin M, Schultz DM, Aimsworth J, Lakshminarayana R, Hellman B, James B, McBeth J, Sanders C, Dixon WG. Engagement and subsequent attrition of daily data entry: Pilot smartphone study of weather, disease severity and physical activity in patients with rheumatoid arthritis. *JMIR mHealth and uHealth* 2017;5(3): e37.

48) Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, Mohr DC. Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study. *J Med Internet Res* 2015;17(7): e175.

49) Blyth FM, March LM, Brnabic AJ, Jorm LR, Williamson M, Cousins MJ. Chronic pain in Australia: a prevalence study. *Pain* 2001;(89): 127–134.