And Now For Today's Pain Forecast: Putting Predictive Models in Patients' Pockets

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health.

2023

Claire L. Little

School of Biological Sciences

Contents

Word Count: 43,548

List of Figures

Abstract

Chronic pain is a common symptom of many long-term health conditions and is a key cause of years lived with disability. However, pain severity is variable and seemingly unpredictable. People with chronic pain report that pain variability creates feelings of uncertainty and impacts daily living. Being able to forecast future pain offers the opportunity to support people living with chronic pain in reducing this pain-related uncertainty.

In this thesis, I aim to forecast pain severity using data from a mobile health study. The specific objectives are to (1) conduct work with patients to identify the desired outcome of a forecast, (2) to conduct cluster analysis of pain severity data to understand common pain patterns, (3) to explore pain variability on an individual level within each cluster, and (4) to forecast movement between clusters.

To prioritise the outcomes of a pain forecast, I conducted a focus group and a survey (Chapter 3). People with chronic pain reported that a pain forecast could be used in planning daily tasks and social events. They prioritised outcomes related to pain flares and fluctuations in pain severity. However, concerns remained around data protection, and anxiety about predicted pain.

To understand common patterns of pain severity, I conducted a cluster analysis of weekly pain trajectories (Chapter 4) and examined transitions between clusters in consecutive weeks. This study reported four clusters representing no/low pain, mild pain, moderate pain, and severe pain. It also showed that two thirds of consecutive weeks were assigned to the same cluster, with movement often to neighbouring clusters.

Substantial within-cluster variability remained. This variability was quantified and associated factors were identified (Chapter 5). Trajectories within the no/low pain cluster were more likely to be stable (no day-to-day fluctuations) than in other clusters. When fluctuations were observed, they were often one-unit (of a possible 4), in all clusters. Fluctuations were associated with pain interference, fatigue, mood, morning stiffness, and participant wellbeing.

To forecast pain, a model was developed for between-cluster movement on consecutive weeks (Chapter 6). Using an elastic net penalty term and different groups of candidate predictors, optimal models for each originating cluster were identified. Common predictors were measures of pain severity, pain interference, morning stiffness, fatigue, dewpoint temperature, and the number of pain conditions.

Future work to forecast pain could explore different outcome measures (e.g., pain variability), use different measures of daily data (e.g., from wearable devices), and examine individual-level models.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

I am incredibly grateful to the team of people who have supported me throughout my PhD journey.

I would like to express my gratitude to my primary supervisor, John McBeth, who has been a constant support throughout my research. His wisdom has helped me to stay focussed and encouraged me to develop as a researcher. Most importantly, he has given me the confidence to persevere, and I appreciate all of the support he has offered.

I would also like to thank my co-supervisors: Will Dixon, Thomas House, and David Schultz. This project has been incredibly interesting at the intersection of different disciplines and I have greatly benefited from the expertise of each of my supervisors. I really appreciate the time taken, especially in commenting on multiple copies of the manuscripts included in this thesis.

In addition, I thank Michelle Briggs for acting as my academic advisor and providing wisdom and support during the tougher parts of the PhD journey.

I extend my gratitude to my colleagues in the Centre for Epidemiology, whose support came in many forms. To Katie, for your expertise in engaging with patients and the public, and for your encouragement to believe in myself. To Belay, for your fantastic statistics help, much of which I seem unable to remove from the whiteboard in our PhD study. To David, for your R support service, and for the many bike rides. To Ilina, for your support in advertising my studies, and for the countless doughnuts. To my fellow PhD students, especially to Danielle and Celina for sharing the ups and downs of the PhD journey and always being a great sounding board.

I'd like to thank my family, especially Mum, Dad and Stephen, for being a constant well of support and always encouraging me to follow my dreams. I really couldn't have done this without knowing that you're always in my corner.

And finally to my rock, Alex. Doing a PhD is an absolute rollercoaster and you've stuck by my side through all the highs and lows. Thank you for everything.

The Author

Claire is a PhD candidate from the Centre for Epidemiology Versus Arthritis. She gained a BSc (Hons) in Mathematics from Durham University in 2016. During her degree, she took optional courses in statistics and mathematical biology. She gained a Postgraduate Diploma in Education (PGDE) from the University of Sheffield in 2017, and later a Masters in Education in 2019.

In 2020, Claire successfully applied for funding through the Centre for Epidemiology Versus Arthritis to conduct the research described in this thesis. She was supported by an inter-disciplinary team of supervisors: John McBeth, Will Dixon, David Schultz and Thomas House.

Funding Awards

- Three-year studentship through the Centre for Epidemiology Versus Arthritis
- Funding from the Centre for Epidemiology Versus Arthritis to undertake Patient and Public Involvement work
- Funding from the Royal Statistical Society to attend the 2022 conference

Conference presentations

- Attendance at IEEE CBMS (Computer Based Medical Systems) 2021
- Attendance at RSC (Research Students Conference) in Statistics and Probability 2021
- Lightning talk at UK-RiME (Research in Musculoskeletal Epidemiology) 2021
- Attendance at RSS (Royal Statistical Society) 2021
- Poster at PGR Showing off your Research 2022
- Talk at DAGS (Doctoral Academy Graduate Society) 2022
- Poster at Versus Arthritis Fellows' Meeting 2022
- Poster at Postgraduate Summer Research Showcase 2022
- Poster at Versus Arthritis Pain Research in the UK 2022
- Lightning talk and poster at RSS (Royal Statistical Society) 2022

Public Engagement

- Presented at Pint of Science, an outreach in Manchester to present research to the general public.

Teaching and organisation

- Graduate Teaching Assistant 2021–2022
- Co-organizer of Centre for Epidemiology Versus Arthritis Summer School 2021 and 2022
- Postgraduate Representative for Centre for Epidemiology Versus Arthritis 2020–2022
- Committee member for DAGS (Doctoral Academy Graduate Society) 2020–2021

1. Introduction

Betty, along with an estimated 15.5 million people in the United Kingdom [1], experiences chronic pain. She is able to discuss the medium-term and long-term prognosis of her symptoms at clinical visits. However, Betty's pain is not stable between consultations and, like many other people living with chronic pain, she becomes frustrated at the uncertainty of her daily symptoms. This uncertainty leads to Betty feeling that she is not in control of her condition, and she often wishes that she could predict what would happen next. This thesis explores the development of a pain forecast to support people like Betty in better understanding their future pain.

1.1 The impact of chronic pain

Chronic pain is defined as persistent pain that lasts for at least three months [2]. It is a feature of many long-term health conditions as both a primary symptom (e.g., fibromyalgia) or as a secondary symptom (e.g., osteoarthritis) [2]. There are an estimated 15.5 million people in the UK who are affected by chronic pain [1], with over 10% of the UK population experiencing pain that is severely disabling [3]. Worldwide, long-term health conditions with chronic pain as a main symptom (low back pain, migraine, neck pain, osteoarthritis) are among the leading causes of years lived with disability [4].

1.1.1 Societal costs of chronic pain

There are substantial economic costs of chronic pain conditions on a societal level. Costs can be attributed to direct costs (e.g., primary and secondary care) and indirect costs (e.g., welfare payments and absence from work) [5]. The NHS report that up to 30% of GP consultations are for musculoskeletal conditions [6], with direct costs in supporting people with musculoskeletal conditions reported to be £5 billion [7]. Further, people with long-term conditions in general are less likely to be employed [8]. There are difficulties in returning to work including managing pain at work, work relationships, and reasonable adjustments [9]. When people with chronic pain are employed, nearly 20 million work days were lost to sickness in 2022 due to musculoskeletal conditions [10], and one third of people who take time off work due to chronic pain remain away from

the workplace after a month [11]. In Europe, the overall cost of chronic pain has been estimated to be 1.5–3% of GDP [5], and the social and healthcare costs increase for individuals with higher pain severity [12].

1.1.2 Individual-level impacts of chronic pain

Chronic pain also impacts individuals' daily living. Broadly these implications include the ability to participate in work, family life and hobbies, as well as the ability to care for family and live independently [1]. For example, a quarter of people with chronic pain report having to change job role, resulting in a loss of earnings [13]. Others report that the physical limitations of their illness requires micromanaging of daily activities [14]. A study among people with chronic pain reported a decreased ability to conduct household chores (54% of respondents), attend social events (48%), and maintain relationships with friends and family (27%) [13]. Further, chronic pain is associated with worse mental health outcomes, including higher levels of anxiety and depression [15, 16]. As a result of these impacts, people with chronic pain report a lower quality of life than the general population [17, 18]. Among people with chronic pain, higher pain intensity is further related to lower health-related quality of life [19–21].

1.1.3 Pain-related uncertainty

One key theme about the pain experience is that it is variable, which leads to feelings of uncertainty around future pain [22]. People with chronic pain have reported that variable pain means that making plans is difficult [14], that social events require risk calculations [22], that uncertainty about future pain increases stress [23], and that uncertainty leads to feelings of not being able to cope [24]. Reducing uncertainty around future pain severity has the potential to support patients and improve their quality of life. As people with chronic pain can become frustrated that their pain seems illogical and that they have no control over their symptoms [22], having clarity about their pain pattern and possible triggers may assist in accepting the nature of their illness and providing a sense of control. Further, a sense of control over chronic pain can improve everyday tasks [1] and learning about when periods of lower pain may occur could assist in planning activities within the home, at work, or socially. There is therefore a need to address pain-related uncertainty to support individuals with chronic pain in their daily living, to reduce stress, and to improve confidence in coping with pain.

1.2 Digital health technologies

Digital health technologies, including wearable technologies and smartphone applications (apps), offer opportunities to monitor and support people with chronic pain, by augmenting visits to primary or secondary care, or by providing support outside of clinical care [25]. One example of the use of digital health technologies to augment clinical care is the opportunity for remote monitoring of pain-related symptoms [26, 27]. This passive monitoring can support clinician-patient conversations by reviewing tracked symptoms over time rather than relying on one-time reporting of symptoms at a clinical visit [28]. An example of providing support outside of clinical care is that digital tools can support people with long-term health conditions through smartphone apps by providing education about their condition, and by supporting treatment regimens [29].

Digital health technologies support different stakeholders. Healthcare professionals have reported the benefits of wearable technology in supporting clinical visits, by tracking objective measures (e.g., of exercise), for identifying high-risk patients, and for supporting self-management of symptoms [30]. Patients have reported that they are motivated to provide data through smartphone apps because it can support their health goals, it can allow tracking of symptoms, and it can support their clinical visits [31].

1.2.1 Patient-generated health data

Digital health technologies support the collection of *patient-generated health data*. These data are any data provided by patients, and include biometric data, questionnaires and surveys, and health history [32]. The collection of patient-generated health data through wearable technologies and mobile apps allows data to be collected in real time and in patients' own environments [33, 34]. Collecting data in this way reduces the burden on participants and minimises recall bias [34].

Patient-generated health data from many participants across multiple time points can form large datasets, requiring sophisticated methods to analyse. *Machine learning* is a strand of artificial intelligence that can be used to analyse large datasets. Machine learning incorporates methods that can be used to extract information, detect patterns, and predict new data [35], especially among datasets containing a large number of predictors [36]. Application of these methods with data from people with chronic pain

have included disease diagnosis, clustering of patient data, predicting disease outcomes and progression, and predicting drug-treatment response [26, 37].

1.2.2 Mobile smartphone applications

Digital health technologies include *mobile health* (mHealth) *apps*, which are smartphone apps that aim to support and improve the health of their users [38]. Mobile apps can be used for the collection of patient-generated health data or to support people with long-term health conditions. In a 2015 study, Krebs and Duncan reported that over half of their US respondents in the general population used a health-related app [39]. There exist a high number of mobile apps that support people with a range of long-term health conditions, including diabetes, hypertension, and arthritis [40].

Dorsey et al. [41] identified a number of benefits in the use of mHealth apps. These include that data can be collected frequently, can engage participants by visualising their data, and can include participants who may otherwise struggle to participate in research (e.g., through their condition, or geographical limitations).

However, the apps purporting to support people with long-term health conditions often have associated risks. Many apps do not use (and sometimes contradict) scientific evidence, they often lack the input of clinicians and patients, and they are often not validated or effective [42]. For example, a number of mental health apps do not use evidence-based treatments [43] and some apps to support contraception promote techniques with a low effectiveness, hence increasing the risk of unintended pregnancies [44]. Further, apps for people with long-term health conditions often provide general information or summarise patient-inputted data, but rarely provide tailored information based on the users desires or needs [40].

The following subsections explore apps that have been designed to support people with chronic pain, and apps that have been used to predict health-related events in other long-term health conditions.

Apps to support people with chronic pain

Mobile apps have been developed to support rheumatic and musculoskeletal conditions. These apps support symptom tracking (including with visualisation), promote physical activity, reduce catastrophizing, and improve patient education [45, 46]. A survey of people with chronic pain reported that around half of respondents used an app to self-

manage their pain [47]. The reasons for use of an app were to track symptoms, communicate with healthcare professionals, and to learn about their disease.

Some apps have been shown to improve pain-related symptoms. For example, apps supporting pain management have been shown to reduce pain severity [48, 49]. Krisjansdottir et al. [50] reported a clinical trial that reduced pain catastrophizing and increased pain acceptance. Although the app required active clinical input, it supported pain management without the need for in-person clinical visits. Meheli et al. [51] developed another app without the need for clinical intervention. This app used a conversational assistant to provide mental health support to individuals with chronic pain. The advice offered by the app was tailored to the user's conversation.

To support people with chronic pain on an individual level, apps have been developed that provide individualised advice. One area that has seen development of such individualised apps is in exercise treatment plans for people with chronic pain. selfBACK is an example of an app that has been developed to provide personalised exercise suggestions for people with low back pain [52]. This app uses case-based reasoning to match new users with previous users and tailors recommendations based on the success of the previous user. A pilot study of MyBehaviourCBP also recommended exercise activities to individuals with chronic pain [53]. In this case, an individual's activities were clustered and recommendations were ranked to maximise those which would be actionable and beneficial. A further personalised exercise program, Well Health, used a multi-layered perceptron artificial neural network to tailor an exercise program for people with neck and back pain using information from a self-report questionnaire [54].

However, there are challenges associated with pain-related apps. The development of pain apps often lacks the input of healthcare professionals or the intended users of the apps during the development process [55]. Further, apps can be inaccessible due to poor functionality, or provide a time burden on their user [47].

Mobile apps to predict health-related events

Health apps have also been used for the prediction of future events. However, few of these apps are evidence-based. As an example, there are several apps that aim to predict the fertile window in women's menstrual cycles, to support women in conceiving. However, of the 73 apps surveyed by Johnson et al. [56] none published their methods,

and only one published their predictive accuracy, which was moderate at 60%. Another review by Moglia et al. [57], identified 108 menstrual tracking apps but only 20 provided accurate calculations, and of these only one cited literature, and only one consulted a professional. Assessing apps is important to ensure their precision, safety, and useability [58] but the lack of published methods means that we cannot rigorously assess these apps, and the lack of scientific backing or consultation with stakeholders reduces the utility of these apps.

Due to a lack of regulation on both the Google and Apple app stores, there are several mobile health apps that use crude methods to predict future results. My Pain Diary [59] claims to support patients in analysing their symptom data and potential triggers of pain. However, the app simply reports visual tracking of previous pain, and provides no information to support identification of triggers. Another app, Dialik [60] claims to provide personalised information about future blood pressure. However, an email from the company revealed that the 'personalised' prediction was based only on the location of the phone (Dialik, personal communication, 30th November 2020). While location is useful to consider weather and pollution factors, there are many other well known risk factors of high blood pressure that should be included in such a model, including age [61].

1.3 Forecasting justification and pipeline

So far, I have shown that there is a need to manage pain-related uncertainty among people with chronic pain, as this can lead to individuals feeling overwhelmed, stressed, and as though they can't participate in social events. I have also shown that digital health technologies enable the collection of patient-generated health data, and that in a few cases, these data have been used within mobile apps to support people with chronic pain, and to predict future events in other long-term health conditions.

The overall aim of my PhD research is to use longitudinal data from an mHealth study to forecast pain severity among people with chronic pain. In this section, I outline the feasibility of forecasting pain, and introduce a pipeline for building a pain forecast. The formal aim and objectives of the thesis build on this pipeline and are provided in section 1.5.

1.3.1 Feasibility of forecasting pain

Forecasting pain severity is theoretically feasible because there are a number of known associations between pain severity and other symptoms. The widely accepted biopsychosocial model acknowledges that pain severity is impacted by biological, psychological and social factors and that these are intertwined [62], meaning that there are many potential predictors of pain. Known associations include genetic, social, psychological, and childhood factors [63]. Evidence indicates that poor sleep quality is associated with increased pain severity on the following day, and that long-term poor sleep is associated with developing chronic pain [64, 65]. Physical activity, including the number of steps per day, is inversely associated with having worse chronic pain [66]. Pain severity can be improved with increased exercise and so exercise is often used as a treatment for chronic pain [67, 68]. Negative mood is associated with worsened pain severity [69], and chronic pain conditions are often comorbid with depressive conditions [70, 71]. Although research into the causal relationships between pain severity and other covariates is limited, it would be possible to exploit these relationships for prediction, using aforementioned covariates as candidate predictors. Further, these covariates (sleep quality, physical activity, mood) can be easily measured using digital health technologies, meaning that they could be useful in a forecasting model that requires self-reported data.

1.3.2 Research Pipeline

To build a pain forecasting model, four stages of research were identified and are shown in Figure 1. These are described briefly in this subsection and explored fully throughout this thesis.

The first stage in the research pipeline was to prioritise outcomes with people with chronic pain. As previously described, many mobile health apps do not involve relevant stakeholders in the development of forecasting models [46]. However, it is important to include stakeholders in research that will directly affect them (in this case by producing a model that would be used by people with chronic pain) [72]. As well as a moral duty in research to include people who will be affected by it, involving relevant stakeholders ensures that the outcomes are relevant and important [72]. Examples of potential outcomes for a pain forecast might include: absolute values of pain severity on the same

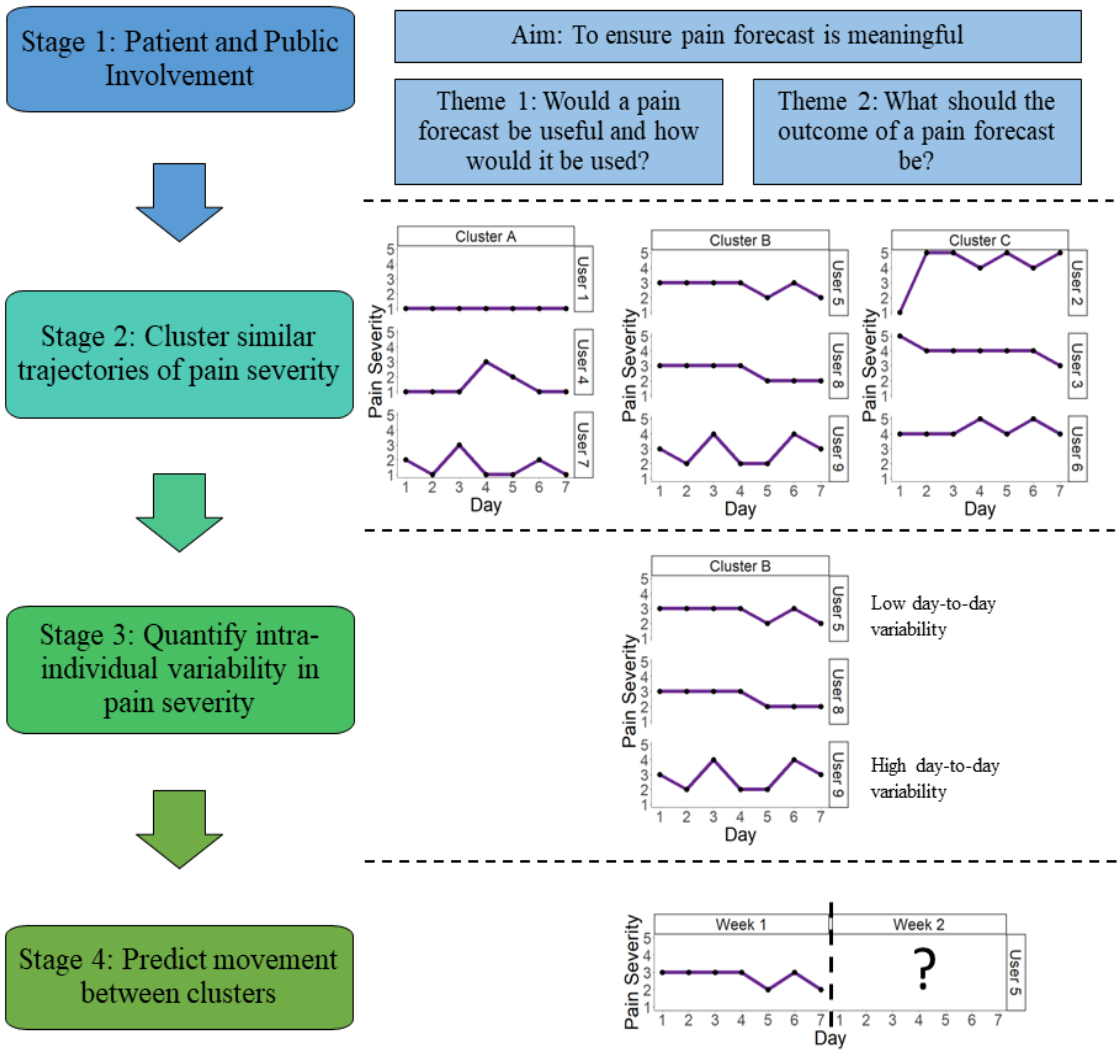


Figure 1: Flowchart of key stages of research

scale as data collection, the occurrence of pain flares (increased pain) or day-to-day fluctuations in pain severity.

The second stage in the research pipeline was to cluster trajectories. It is possible that group-level associations are masked by population-level analysis and so it is important to identify meaningful groups of data, within which to conduct future analysis. In his paper, Simpson [73] showed that it would could be possible for variables *A*, *B* and *C* to fulfil the equation

$$P(A \cap B|C) > P(A|C)P(B|C)$$

for all values of *C*, even if *A* and *B* were independent when *C* was not included. It is even possible that the relationship between *A* and *B* could be inverted depending on the presence of *C*. To illustrate this paradox, simulated data with two continuous variables *A* and *B* are shown in Figure 2. The left panel examines the relationship between *A* and *B* and a positive relationship is seen ($r = 0.61$). More formally, a linear regression provides a positive slope with $p < 0.001$. The right panel highlights a third, categorical variable *C*, highlighted by different shapes on the graph. Within each factor of *C*, linear regression returns a negative value of the slope with $p < 0.001$ for each factor.

These simulated data highlight Simpson's Paradox: that relationships between variables can be reversed when a latent categorical variable is included in a model. Another example is reported by Haigh [74]. Two drugs, *D* and *E*, could be administered, and *D* may produce fewer side effects to the population as a whole, but drug *E* may produce fewer side effects to both men and women. Simpson's Paradox therefore highlights the importance of considering subgroups of data, as population-level associations may mask subgroup-level associations.

Subgroups of pain data are important as studies have shown that individual pain trajectories often do not follow the population average, and therefore more nuanced descriptions of the pain experience are required. A review by Maixner et al. [75] highlighted that there is heterogeneity in the risk factors for chronic pain across different conditions, and that pain across conditions may have clusters of individuals. Further, previous work has identified factors that are associated with membership to a certain cluster. For example, Kongsted et al. [76] summarised that studies have found factors including participation in work, mood, expectations of recovery and comorbidity to be associated with membership to different clusters of low back pain. It is therefore possible that there are clusters of pain severity that may have different associations with other factors described in Section 1.4.1, and that could be exploited in a predictive model.

The third stage of the research pipeline was to describe variability in the trajectories identified in Stage 2. People with chronic pain have described that their pain is changeable, varying within- and between-days [77]. It is this changeable pain that has been reported to impact forward-planning and attending social events [78]. Increased pain variability has also been associated with outcomes such as increased opioid use

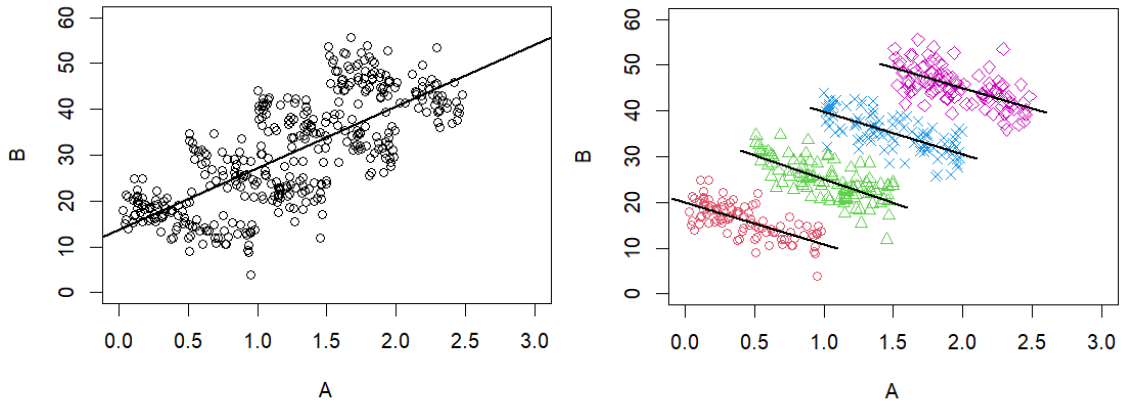


Figure 2: Demonstration of Simpson's Paradox using simulated data. Continuous variables *A* and *B* are simulated. In the left panel, a positive association between *A* and *B* is seen. In the right panel, colours represent categorical variable *C*, and in each subgroup a negative association between *A* and *B* is found.

[79], although it is associated with improved response to treatment [80] possibly due to the possibility of movement to a lower pain state. I hypothesise that better understanding within-cluster variability may identify factors associated with this variability and may support forecasting movement between clusters.

The fourth stage of research was to forecast movement between clusters. Forecasting this movement would act as a first step towards summarising pain worsening (into a more severe pain cluster), pain recovery (into a less severe cluster), or pain stability (remaining in the same cluster). Forecasting between-cluster movement would require minimal data collection from participants (the length of one trajectory) before predictions could be provided. Further, modelling between-cluster movement would use multiple trajectories from each cluster and hence improve the predictive power achieved.

To model between-cluster movement from an *origin* cluster to a *destination* cluster, separate models for each origin cluster can be developed. The outcome measure in this case is the destination cluster. As cluster membership is a nominal variable, multinomial models can be used [81]. Predictors used in these models should be interpretable so that potential triggers of pain can be identified. Summary statistics of patient-generated

health data offer interpretable candidate predictors. *Feature selection* can be conducted while developing models to select important predictors and deselect features of lower importance.

1.4 Data

1.4.1 Data source

Data used throughout this thesis are from a UK-based population based mHealth study, Cloudy with a Chance of Pain. Data collection and summaries of other analyses using these data are described in this section.

Study recruitment

Cloudy with a Chance of Pain was a longitudinal observational study which collected data between January 2016 and April 2017 [82]. Individuals with chronic pain conditions were recruited following advertisements on television, radio, and social media. Interested participants were recruited through a smartphone app. On the app, electronic consent for the study and demographic data were collected. These demographic data were year of birth (entered as free text), sex (male or female), chronic pain condition (selected from pre-defined responses e.g., osteoarthritis, fibromyalgia/chronic widespread pain), site of pain (selected from pre-defined responses e.g., knee, hands, multisite), medication (selected from pre-defined responses e.g., paracetamol, nonsteroidal anti-inflammatory drugs, synthetic disease modifying antirheumatic drugs), belief in the likelihood of a relationship between pain and the weather (entered on an 11-point numeric rating scale, 0 = not at all likely, 10 = extremely likely). There were 13,207 participants who downloaded the app and provided electronic consent and demographic information [82].

Collection of data

Participants were asked to report ten factors daily for at least six months. These patient-generated data were reported on a five-point scale by dragging a motif (Fig 1 [82], reprinted here as Figure 3). Individual items were (Figure 3): pain severity (1 = no pain, 5 = very severe pain), sleep quality (1 = very poor, 5 = very good), feeling tired (1 = not at all tired, 5 = extremely tired), fatigue (1 = no fatigue, 5 = very severe fatigue), time spent outside (1 = none of the day, 5 = all of the day), activity (1 = no exercise, 5 = 30+

Data item	Question stem	Anchor - bottom	Anchor - top
Pain severity	How severe was your pain today?	No pain	Very severe pain
Fatigue	How severe was your fatigue today?	No fatigue	Very severe fatigue
Morning stiffness	How stiff did you feel on waking this morning?	No stiffness	Very severely stiff
Impact of pain	How much has your pain interfered with your activities today?	Not at all	Very much
Sleep quality	How was your sleep quality last night?	Very poor	Very good
Time spent outside	How much time have you spent outside today?	None of the day	All of the day
Waking up feeling tired	How tired did you feel when you woke up this morning?	Not at all tired	Extremely tired
Physical activity	How long have you exercised today?	No exercise	30+ minutes of strenuous exercise
Mood	How has your mood been today?	Depressed	Very happy
Wellbeing	How well did you feel today?	Very unwell	Very well

Figure 3 reprinted from Figure 1 in Dixon et al. (2019) How the weather affects the pain of citizen scientists using a smartphone app. NPJ Digit Med 2:105. Left panel shows the motif used by participants to report symptoms. Right panel shows the questions and anchors for each data item.

minute strenuous activity), pain interference (1 = not at all, 5 = very much), mood (1 = depressed, 5 = very happy), wellbeing (1 = very unwell, 5 = very well), morning stiffness (1 = no stiffness, 5 = very severe stiffness).

When patient-generated data were provided, the date, and time were passively collected. Each hour, the location of the smartphone through the global positioning system (GPS) was recorded, and used to locate the closest weather station [83]. Weather variables (windspeed, temperature, dewpoint temperature, pressure, and relative humidity) were collected from these local weather stations and means of each weather variable experienced by an individual throughout the day were calculated. The calculation of weather conditions in this way accounted for movement of participants [83].

Data types

Considering both the patient-generated data and passive data collected during this study, there were four types of data. These were: binary data (e.g., sex), categorical data (e.g., diagnosis), ordinal data (e.g., pain severity), and continuous data (e.g., windspeed). Some of these data were time invariant for each participant (e.g., sex and diagnosis), while others were time varying and formed longitudinal data.

When data were collected longitudinally, some participants provided data on just one day, while others provided complete data for over six months. Others provided some periods (e.g., a week or a month) of complete data and other periods with missing data. To account for the different amounts of data reported, I define a *weekly pain trajectory* in this thesis as seven days of complete pain severity data provided by a single participant. This weekly pain trajectory begins on a Monday and ends on a Sunday although assumptions about this definition will be tested. The weekly trajectories require only one week of data from participants, while allowing for multiple weeks of trajectories from long-term participants. Weekly trajectories introduce a social meaning to the data, for instance by weekends being in the same location in the trajectory.

1.4.2 Previous analyses

Previous analyses have examined various relationships within the complex data from Cloudy with a Chance of Pain. On a population level, Dixon et al. [82] used a case-crossover analysis to compare weather on days when participants experienced a pain event (an increase of 2 or more in pain severity across consecutive days) with a control day in the same month (a day when participants were eligible to have a pain event but did not). They found that the odds of a pain event increased with an increase in wind speed and relative humidity or a decrease in atmospheric pressure. Schultz et al. [84] examined days with a high fraction of participants recording pain events (defined by a +1 increase in pain over consecutive days). They found that these days had a significant relationship to the temperature, dewpoint temperature, pressure and windspeed. To extend the understanding of these data, I conducted analysis with a group of other researchers to explore the impact of spending time inside or outside buildings on the relationship between the weather and pain and found that time spent outside partially mediates the weather-pain relationship. A manuscript describing this work has been

submitted to *Weather, Climate and Society*. I am the first author of this submission and the manuscript is provided in Appendix A.

Yimer et al. [85] explored individual relationships between pain events and weather variables, using a multilevel ordinal probit model. They reported heterogeneity between these relationships. For example, one tenth of participants had significant associations between pain events and the temperature, but the direction of associated differed for individuals. Das et al. [86] used Markov processes on dichotomised mood and pain severity data. They reported four clusters based on the transition between mood-pain states. These two studies identified that there are individual differences in the relationships between exposure data and pain events.

1.4.3 Challenges in analysing data

Despite the benefits of using digital health technologies to support the collection of patient-generated health data, there remain challenges associated with analysing these data. These challenges are common across many studies of patient-generated health data, including Cloudy with a Chance of Pain.

First, some data are collected on ordinal scales. An example of ordinal data in health data are the responses to questions on a Likert scale such as “Strongly disagree”, “Disagree”, “Neither agree or disagree”, “Agree”, “Strongly agree” [87]. In Cloudy with a Chance of Pain, ordinal data included the self-reported data such as pain severity. Common methods of analysis, such as linear regression cannot be directly used with data collection on the ordinal scale as they do not possess metric qualities [87]. For example, assuming metric properties with 1 = “Strongly disagree” and 5 = “Strongly agree” suggests an equal difference between “disagree” and both “strongly disagree” and “neither agree or disagree”, and also that “disagree” is half the value of “agree”. Methods for analysing ordinal data should therefore be considered [87]. For example, ordinal probit models assume an underlying latent normal distribution and assign probabilities to each ordinal value using **intervals** of probability density between values. On the contrary, a metric model assigns probabilities using the value of probability density at exact values [87].

Second, the alignment of data should be considered. Data collected multiple times from the same participant following an explicit event, such as an operation or the start of a treatment, can be aligned across different participants. That is, time since the event can

be used to compare data between participants. However, data collected in participant's own environments in an observational study may not have such an event, other than the beginning of a study. In this case, time since the start of the study is not an anchor to compare different participants, and trajectories resulting from these data are *time-invariant* [88]. For example, consider two trajectories of pain severity that record pain severity as (1,5,1,5,1,5,1) and (5,1,5,1,5,1,5). These two trajectories are very similar because they alternate between pain scores of 1 and 5 throughout the week. However, their starting values are different which result in matched days always having different pain severity between these weeks. The potential for time-invariance in data should therefore be considered when analysing patient-generated health data.

Third, analysing patient-generated data requires the output to be explainable. In healthcare, understanding the reason for algorithmic output is particularly important [89], especially in high-stakes situations. For example, clinicians ethically require an understanding for the recommendations of algorithms that may recommend performing an operation over other actions [90]. In other situations, explanations of algorithms support building trust with relevant stakeholders through their transparency [89]. However, many algorithms used in clinical practice are so-called 'black box' algorithms, where output measures such as coefficients are not explainable, either because they are too complicated or they use proprietary algorithms [90]. Attempts to better understand black box outputs have led to an increase in 'explainable AI'. However, black box algorithms do not necessarily perform better than simpler (e.g., linear or logistic regression) methods, and 'explainable AI' methods are not successful all of the time, leading to an added layer of potential misinterpretation of the output [90].

Fourth, machine learning is susceptible to overfitting [35]. Overfitting occurs when a model follows training data too closely but would not perform well on unseen (test) data [91]. Overfitting can occur if the training model erroneously identifies associations between covariates and the outcome variable by chance but the association is not observed in the test data [91]. A number of methods can minimise the impact of overfitting. For example, one can retain a test set that is separate to the training data. Alternatively, *k*-fold cross-validation splits the data into *k* equally sized groups, trains the data on *k*-1 sets, and tests on the remaining set. This process is repeated *k* times, leaving out a different set each time.

1.5 Aims and Objectives

The aim of this thesis was to forecast pain severity among people with chronic pain, using data from a mobile health study. To achieve this aim, the framework presented in Figure 1 was followed. The objectives were to (1) identify an outcome for the pain forecast by conducting work with people living with chronic pain, (2) identify common clusters of pain severity using patient-generated health data, (3) explore variability of individual trajectories within these clusters, and (4) forecast movement between clusters and identify predictors of this movement.

1.6 Note on journal format of thesis structure

This thesis is presented in the journal format. The work presented is likely to be of interest to a wider audience and therefore I felt that publishing work relating to this thesis was important. The articles presented would be of interest to different audiences, including pain specialists, those who work directly with patients, and statisticians/data scientists. Three of the four manuscripts provided as main chapters have been submitted for peer review and the journals were selected based on the target audience for each paper.

The remaining chapters are structured as follows. Chapter 2 provides a discussion around methods previously used for clustering, defining variability, and forecasting. This chapter is not presented in the journal format. Chapters 3 – 6 follow the journal format. They address the four objectives of this thesis in turn. At the start of each chapter, a brief introduction is provided. This is followed by a manuscript which has been (or will be) submitted to a journal for publication. Chapter 7 provides a discussion and conclusions of the overall thesis, combining all chapters together. Further work, including another submitted piece of work in which I was first author and supplementary materials for each paper, are provided as Appendices.

2. Literature Searches

The pipeline outlined in Section 1.4 described four stages of research. Stages 2–4 represent different machine learning approaches to understanding a complex dataset. The individual stages described (clustering, exploring variability, forecasting cluster movement) build on previous literature. Therefore, to identify previously described methods and results in these areas, I report on literature on clustering pain trajectories (Section 2.2), pain variability (Section 2.3) and forecasting outcomes in long-term health conditions (Section 2.4).

2.1 Clusters of longitudinal pain severity data

To explore studies related to Stage 2 (Figure 1), I conducted a systematic search of literature. The aim was to understand previous research in clustering pain trajectories. The specific objectives were to (1) identify methods used for clustering pain trajectories, and (2) summarise pain clusters that are commonly reported. Methods used in clustering trajectories are not unique to the pain field, but I focus here on pain severity as an outcome due to the vast literature in this field.

Search methods

The inclusion criteria were that:

- Study participants were humans with a chronic pain condition.
- The study was longitudinal, collecting a pain outcome on at least three occasions.
- The paper reported new clusters of pain severity that were identified by clustering methods.
- The paper was written in English.

Exclusion criteria were:

- Study participants had a cancer-related chronic pain as a primary condition.
- Study participants were healthy or were subject to simulated pain.
- Study participants were animals.

- Clusters were pre-defined, including papers that reported further analysis of previously identified clusters.
- Clusters were focussed on other symptoms or medication.
- The paper was a case report, systematic review, study protocol, commentary or editorial.
- The paper was unavailable in English.

To identify studies that used pain as a primary symptom, "pain" was a required term in the title. To identify studies that considered clusters of longitudinal data, preliminary scanning of the literature identified key terms that were used in titles of relevant papers. On 2nd December 2020, I conducted the following search on PubMed, and updated this search throughout my PhD, with the final update on 26th June 2023:

1. pain[Title]
2. temporal[Title] OR longitudinal[Title] OR time[Title]
3. cluster*[Title] OR subgroup*[Title] OR class*[Title]
4. (2) AND (3)
5. trajector*[Title] OR course[Title] OR pattern[Title] OR progress*[Title]
6. (4) OR (5)
7. (1) AND (6)

The final search was (7), with further filters for "Human" and "English". This search returned 1314 papers which were assessed against the exclusion criteria above. Of these, 907 were excluded due to the title, 5 were not accessible via the University of Manchester library, 204 were rejected after reading the abstract, and 91 were rejected after reading the full paper. Figure 4 reports the reasons for exclusion at each stage. The most common reason for exclusion when reviewing the titles and abstracts were that search terms such as "pattern" or "course" were unrelated to symptom data (pain or otherwise). Examples of this included "gait pattern" or "educational course on pain".

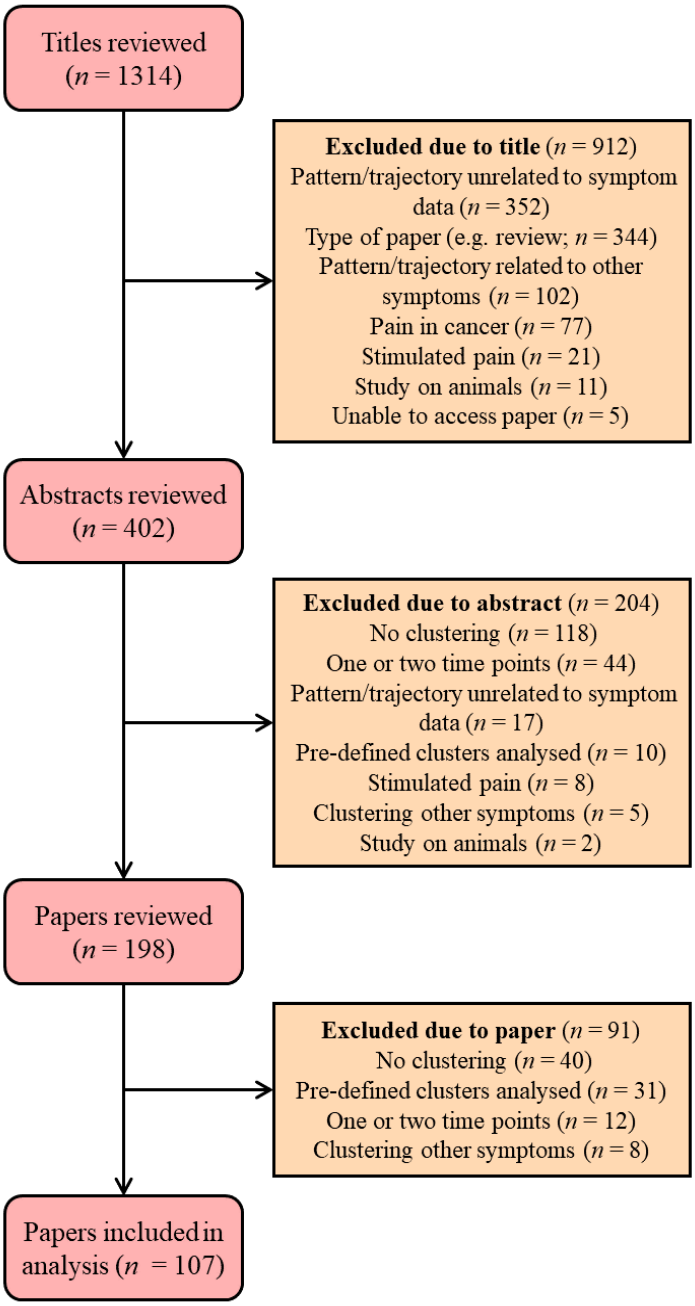


Figure 4: Papers analysed in systematic search. Reasons for exclusion are listed following examination of the title, abstract, and paper.

Search results

Of the 107 papers included in the analysis, 108 studies were identified. Included studies were screened for information about the pain condition and data source used, the number of participants, the mean age of participants, the percentage of females in the

study, the pain score used as an outcome, the frequency of collection of the pain score, the statistical method used for clustering pain scores, the method(s) used for dealing with missing data, the criteria used for comparing models with different numbers of clusters, the number of clusters identified, and the description of clusters identified. Due to the large number of results, the full results are presented in Appendix B. Papers are organised primarily by the duration of data collection, secondarily by the number of data points collected within that time, and further by surname of the first author, and by year of publication if required.

Key features of the table support an understanding of the state of the literature and are summarised in this section. First, the results table reports the year of publication. Figure 5 illustrates these years, highlighting a growing interest in clustering pain trajectories in recent years. Second, the pain condition(s) are listed and summarised in Figure 6. Studies focussed on one chronic pain condition (n = 53), multiple chronic pain conditions (n = 4), postoperative or acute pain (n = 27), or more general chronic or musculoskeletal pain (n = 24). Third, the table and Figure 7 report the frequency and length of data collection. A substantial number (n = 37) of studies collected data for at least five years, and a further 33 collected data for at least one year. However, data collected in these longer studies is often very sparse. Only 24 studies that lasted for at least a week collected ten or more data points in that time, and only 7 studies in total had more than 30 data points per participant.

Methods used in clustering pain trajectories

The first objective of this systematic search was to identify methods previously used to cluster pain trajectories. The table in Appendix B and Figure 8 report the method used in each study. The statistical methods were latent class growth analysis (also called latent class analysis, n = 48), group-based trajectory modelling (n = 27), growth mixture models (n = 23), k-means clustering (n = 7), and hierarchical clustering (n = 3). An overview of these methods is provided below.

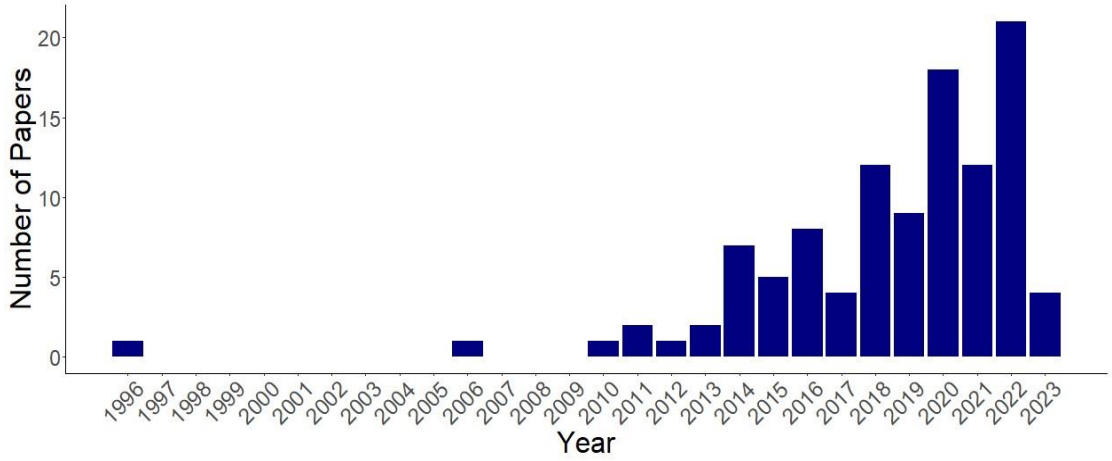


Figure 5: Number of papers by year of publication in systematic search of clusters of pain severity.

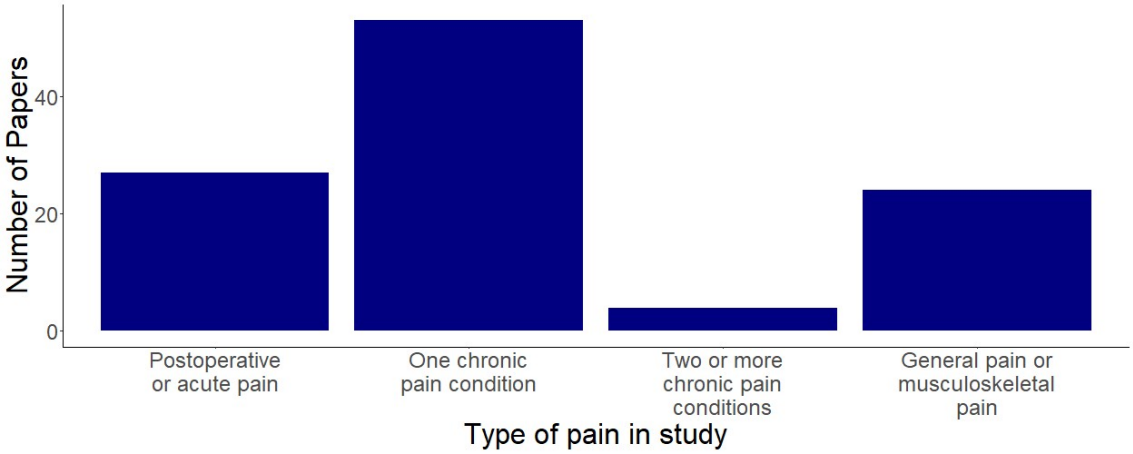


Figure 6: Type of pain studied in systematic search of clusters of pain severity.

The *k*-means clustering algorithm is an example of an expectation-maximisation (EM) method, so named because the algorithms alternate between these two components. The *k*-means clustering algorithm is an unsupervised method used to partition data into a number of clusters, *k*, which is decided a priori. To partition data, the method minimises variation within clusters [91]. First, each point is randomly assigned to a cluster. Second, the centre of each cluster is defined as the mean of the data. Third, each data

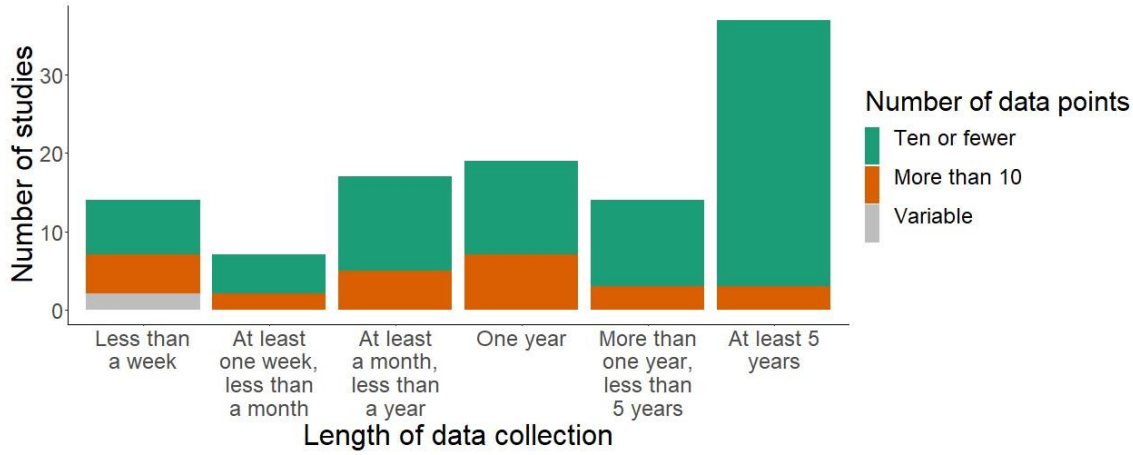


Figure 7: Data collection in studies from systematic search of pain clusters. Bars show the length of data collection and colours show the number of data points collected for each participant in the study.

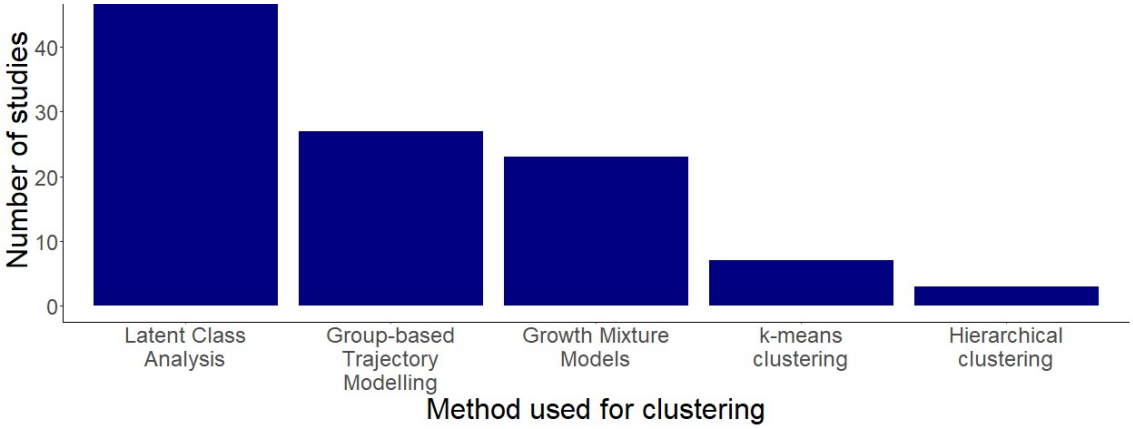


Figure 8: Clustering methods identified in systematic search of clustering pain trajectories.

point is reassigned to the cluster for which its Euclidean distance to the centre is the smallest. Steps 2 and 3 are repeated until no changes are observed [91]. *KmL* is an R package which contains an adaptation of *k*-means clustering for longitudinal data. In this adaption, the distance between individuals *i* and *j* for each time point $t \in [1, T]$ and for each pain outcome *y* is $\sqrt{\sum_{t=1}^{T}(y_{it} - y_{jt})^2}$ [92]. By minimising the squared distance between individuals at each time point, the resulting clusters contain trajectories that are

as close as possible to each other at each time point. A key benefit of the *KmL* package is that the algorithm is non-parametric meaning that assumptions about the form of the data do not need to be made [92]. Within the context of pain trajectories, this also means that assumptions about the shape of the trajectories are not required. However, *k*-means clustering suffers from its reliance on the Euclidean distance, and it does not consider the density of the clusters [93]. For example, in two dimensions the Euclidean distance may fail to identify oblong shapes. Further, the use of the mean in the algorithm makes it particularly susceptible to outliers [93]. Solutions to both of these have been suggested, including using the Manhattan distance [92] within a *k*-medoids clustering algorithm [94]. The *k*-medoids algorithm is within the same family of methods as *k*-means clustering but a datapoint from each cluster is used as the centre, which makes the clusters more interpretable. The algorithm is more flexible in its use of distance measures than the *k*-means algorithm, and specifically can be used in conjunction with the Manhattan distance. Using the Manhattan distance reduces the dependence on outliers as it does not square (potentially large) distances between a data point and the centre of the cluster.

Another method used for clustering data in the literature is hierarchical clustering, which aims to build a hierarchy of the trajectories [95]. There are two overarching methods for completing hierarchical clustering: agglomerative and divisive [95]. Agglomerative clustering treats each individual as a cluster and groups the two most similar clusters until all data is in the same cluster. On the other hand, divisive clustering operates in the opposite order: all data start in one cluster and are broken down until each individual belongs to a unique cluster. The results of both methods are usually displayed in a dendrogram and a number of clusters chosen visually [96].

The remaining methods fall within the category of structural equation models. Latent class analysis assumes a number of distinct clusters and zero variance within clusters. In other words, the analysis assumes that within each cluster, there is a ‘true’ latent (unseen) curve that all trajectories within the cluster follow and that any within-cluster variation is due to random noise [97]. Growth mixture modelling allows for heterogeneity within groups [98] by allowing a non-zero variance structure. Group-based trajectory modelling also allows for heterogeneity, and goes further by assuming a spectrum of trajectories rather than assuming clusters and aims simply to cluster the most similar ones together [99]. Of the structural equation models discussed here, group

based trajectory modelling makes the fewest assumptions about the nature of the latent trajectories. All of these SEM make assumptions about the form of the data which may be unsuitable for clustering patient generated health data. First, they examine growth in trajectories, which may be unsuitable to describe within-week variation. Second, the methods use linear or quadratic trajectories. Ordinal data does not possess metric properties such as a pain score "4" being twice as bad as "2", and therefore linear increases in pain cannot be assumed to be meaningful. Validated instruments with many values (e.g., 101-point VAS) use linear increases due to the granularity of the scale, but linear methods would be unsuitable for a 5-point ordinal scale.

Pain clusters identified by studies

The second objective of this systematic search was to describe clusters identified through these methods. The number of clusters reported ranged from 1 to 12, with 106/108 studies reporting 2–6 clusters. Clusters were defined by the severity of pain (e.g., mild pain, severe pain), or by long-term changes in pain (e.g., fluctuating pain, decreasing pain).

This systematic search provided an overview of the methods and clusters identified by previous studies. One limitation was that only one database was searched, and by only one author. It is possible that other terms may have returned further papers in the field, and indeed some papers were erroneously excluded by PubMed when filtering human studies. However, this search provided an understanding of the field, identified key methods used in clustering pain trajectories, and summarised the clusters previously defined.

2.2 Measures of pain variability

To understand current research in daily pain variability, I conducted a systematic search on PubMed. The objectives of this search were to (1) understand methods used for analysing pain variability, and (2) identify covariates associated with pain variability. The search was restricted to pain as other symptoms (e.g., blood pressure) use measures on a continuous scale and so different indices of variability can be used (e.g., standard deviation).

Search terms

The systematic search aimed to identify papers that had measured variability in pain severity using at least daily data. On 13th July 2023, I searched PubMed with the search terms: (pain[Title]) AND ((variab*[Title]) OR (fluctuati*[Title]) OR (vary*[Title]) OR (chang*[Title])) AND ((day[Title]) OR (daily[Title])). This search returned 40 papers. Of these papers, 15 explored variability in other symptoms, 6 did not collect daily data, 1 did not provide enough information about variability measures, 1 was not written in English, and 1 explored data on mice. Of the remaining 16 papers, 17 studies were reported. These studies are summarised in Appendix C, ordered by the duration of data collection, the frequency of data collection and author's surname.

Search results

The papers were written between 2009 and 2023. Nine of the studies used daily data, while eight used data collected multiple times a day. The study durations ranged from one day to 165 days (with 15/17 lasting less than a month). Most studies explored variability implicitly within models. These models were multilevel models, linear mixed models, hierarchical linear models, and dispersion models. Other papers (n = 5) used individual standard deviation or variance as a measure of pain severity. These measures calculated the standard deviation/variance within a single trajectory. This measure of pain is inappropriate for ordinal data as it assumes that data collected are continuous, and on a metric scale. However, work throughout this thesis will use ordinal data and hence cannot use the individual standard deviation as an appropriate measure. Two papers used other methods for exploring variation. One used the range of pain scores (as the difference between the minimum and maximum reported scores). The other deemed a change of 33% on the rating scale as clinically significant and representing useful variability.

The papers reported that variability measures were associated with BMI, site of pain, pain condition, medication, age, race, response to treatment, genotypes, depression, happiness, frustration, and physical activity. Days with higher pain severity (above an individual's average) were associated with fatigue, depressive symptoms, cognitive function, poor sleep, time spent asleep, daily catastrophizing, self-efficacy, negative affect, occupational load, physical activity, and disability. However, some studies

reported that there were no associations between pain variability and explored predictors [100, 101].

Other literature exploring pain variability

Although not returned by the systematic search, an important paper in the field was written by Mun et al. [102], who report four measures for summarising individual pain variability. The first is individual standard deviation (as described above), but the authors admit that this does not capture temporal characteristics. For example, the trajectories reporting pain (1,5,1,5,1) and (1,1,1,5,5) would return the same standard deviation, but the experience of pain fluctuations are different. The second measure is autocorrelation which reports the correlation between current pain severity and lagged pain severity. A limitation of this measure is that it is dependent on the choice of lag. The third measure is the mean square of successive difference which summarises both the temporal feature of autocorrelation and the severity of the individual standard deviation. The measure is calculated by squaring changes in pain severity and calculating the mean of these values. However, this measure is difficult to interpret and is invalid when the pain scale is not metric. The final measure is the probability of acute change. This measure is calculated as the proportion of changes greater than a set value. A limitation is the requirement for the choice of the cutpoint. Due to the limitations of these measures, it is recommended by the authors that multiple measures provide broader impression of the variability within pain trajectories.

2.3 Forecasting methods

The fourth stage of the research pipeline is forecasting movement between pain clusters (Figure 1). To describe relevant literature in this field I identified studies forecasting outcomes in pain-related health conditions, and studies forecasting short-term symptoms in long-term health conditions. When reviewing these studies, I aimed to identify (1) outcome measures, (2) forecasting methods, and (3) covariates associated with the outcome measure.

Forecasting outcomes in pain-related health conditions

Forecasting outcomes in pain-related health conditions have been reported on a population level and on a subgroup level. This thesis is focussed on cluster-level forecasting, but I provide a brief overview of each type of forecasting here.

There is a vast literature to develop and validate prognostic models among people with chronic pain. These prognostic models are most frequently at a population-level. For example, among osteoarthritis patients, a meta-analysis identified 52 studies predicting deterioration of pain, deterioration of physical functioning, or the course of physical functioning [103]. In an umbrella review focussed on neck and low back pain, 41 reviews were analysed to identify common prognostic factors for pain, disability, and return to work [104]. As these were reviews, neither paper reported the methods used for prediction within the studies. Identified prognostic factors were higher pain at baseline, knee symptoms, and depressive symptoms among osteoarthritis patients [103], and disability, mental health, pain intensity, pain severity, coping, expectation of outcome, and fear-avoidance among participants with neck or low back pain [104].

Predictive models have also been developed to identify cluster membership. In one study, clusters of pain volatility among participants reporting any pain were identified using the *k*-means clustering algorithm [105]. Two clusters representing low volatility and high volatility were described and used as the outcome measure of predictive models. Four different methods were used to predict cluster membership and achieved good accuracy: logistic regression with a ridge regression penalty (78.1%), logistic regression with a LASSO penalty (79.0%), random forests (79.0%), and support vector machines (77.6%). It should be noted that these accuracies were for an unbalanced dataset and prediction of membership to a low volatility cluster (96.1% – 99.8% accuracy) was much better than prediction to a high volatility cluster (0.6% – 18.1%). There were 130 predictors in the study although identifying key predictors of cluster movement was acknowledged as an avenue of future work. In another study, clusters of pain impact among older adults were identified using latent transition analysis [106]. These clusters were (1) no/low pain, (2) pain presence but low-moderate pain impact, and (3) pain presence and high pain impact. The outcome measure was cluster membership. Covariates of the transitions were identified by a model-based approach that incorporated the covariates into the latent transition analysis model. In this study, physical activity was identified as a covariate associated with transitioning between clusters.

Forecasting outcomes in other long-term health conditions

There have been examples of using patient-generated data in predictive models. For example, one paper studied the outcome variable of day-to-day changes in hayfever symptom scores [107]. A logistic regression was used to forecast a binary outcome, and a proportional odds logistic regression was used for individual forecasts of scores on a 5-point scale. Candidate predictors were added to the until predictive performance remained constant. Key predictors in the logistic model were medication intake and pollen concentrations reported on the same day and the previous day.

Another study also used hayfever symptoms as an outcome variable [108]. In this study, different models were tested: mutli-layer perceptron (a neural network algorithm), *k*-nearest neighbours, model tree, multiple linear regression, and 'last year repeat now' (acting as a baseline model). Models were trained for each individual user and the authors employed cross-validation, which splits the data into *k* subsets, trains the model on each combination of *k*-1 subsets, and tests on the *k*th in each case. Candidate predictors for the model were ranked using the correlation coefficient with the outcome variable. The authors report good results for individualised forecasts when >100 data points were provided by an individual. However, when <30 data points were provided by individuals, the model performance approached random chance. There were not significant differences in the results produced by each non-baseline model.

A final example used self-reported data from participants with chronic obstructive pulmonary disease (COPD). Symptom data on a 4-point scale were used to create a binary outcome measure representing exacerbations of symptoms. Models used for prediction were a logistic regression with regularization, and a random forest classifier. The authors used cross-validation in model development. The results showed that the random forest classifier performed significantly better than the logistic regression model, possibly due to the inclusion of interaction terms among the predictors. Important predictors were mean summary measures of previous disease and symptom scores.

Implications

Data collection among studies developing predictive models for pain-related conditions is often sparse. For example, the mean follow-up for prognostic models among participants with osteoarthritis ranged from 6 months to 8 years [103]. Transitions

between clusters of pain impact were identified using data collected two years apart [106]. It is therefore unclear whether the predictors identified could be used for a short-term forecast. On the other hand, there are examples of predictive models that have used daily symptom data in other long-term health conditions (here, hayfever and COPD).

Some of the studies cited used logistic regression or similar methods for data prediction. In the example predicting events in COPD, the random forest classifier performed significantly better than logistic regression. However, it is not always true that more complicated models provide better predictive performance [90], and this is seen in forecasting pain volatility [105]. Further, simpler models such as logistic regression can provide more interpretable results [90] than methods such as random forests or support vector machines. It may therefore be preferable to use simpler models to investigate pain forecasts.

2.4 Discussion

Machine learning techniques (clustering, exploring variability, forecasting) will be employed during this thesis (Figure 1). The present chapter aimed to identify previous studies that had used these techniques and to summarise the methods and results from these studies.

Section 2.1 identified methods used for clustering pain trajectories and summarised commonly reported pain clusters. The most frequently used family of methods were structural equation models, but these have the limitation that the form of the data (e.g., quadratic) must be assumed, and that the models assume the data is metric. These assumptions may be valid for long-term trajectories (e.g., across multiple years) but there is no evidence that the assumptions would be valid within weekly trajectories. Another identified method was the *k*-means algorithm, and while this is limited by its assumption that data are metric, other algorithms in the same family (e.g., *k*-medoids) can use other distance measures which may be a pragmatic alternative. The search identified that studies frequently report between 2–6 clusters of pain severity, although 2 studies (out of 108) reported 12 clusters. These clusters are defined by the severity of pain or by long-term changes in pain severity. Therefore, it would be expected that weekly trajectories could be described by 2–6 clusters of pain severity, but Chapter 4 of this thesis will investigate whether weekly trajectories follow similar patterns to other studies.

Section 2.2 identified methods used for quantifying pain variability and covariates associated with such variability. Most studies explored variability implicitly in a model (e.g., multilevel models) but some studies used explicit measures of variability. The most common explicit measure of pain variability was the standard deviation within each trajectory. However, standard deviation is not a suitable measure for ordinal data. Alternative measures of variability have been suggested which are more suitable for summarising the variability of ordinal data. Limitations remain for each measure, and so a combination of these measures will be used in this thesis. Studies in Section 2.2 also identified covariates associated with pain variability. However, some studies identified that there no associations between tested covariates and changes in pain severity, and there is hence a need to investigate these associations further. Many of the identified covariates were collected in Cloudy with a Chance of Pain (e.g., sleep quality) and therefore Chapter 5 can explore whether these covariates are associated with daily changes in pain severity among patient-generated health data.

Finally, studies using predictive models were reported in Section 2.3. Predictive models have been used in a range of healthcare settings, including in chronic pain research. Many prognostic models have used sparsely collected data to inform outcomes of chronic pain. There have also been a limited number of studies investigating movement between pain clusters. Methods used in studies of other long-term health conditions were reported in Section 2.3. Logistic regression was a common method and was sometimes not outperformed by more complex machine learning models (e.g., neural network). Further, models with a linear relationship between predictors were acknowledged as being more interpretable than alternatives. Therefore, a logistic regression model with linear predictors is a pragmatic choice for a predictive model and will be used in Chapter 6.

The literature searches in this chapter were not anticipated to be exhaustive and were therefore limited in a number of ways. First, the limited search terms in the systematic searches may have failed to identify further relevant studies. Second, the searches were only conducted by one author and were therefore at risk of author bias in the inclusion or exclusion of studies. Third, the searches did not critically appraise studies and hence may have included weak studies. However, the searches provided an overview of the methods and results used in relevant studies. As a result, they provide a background to the literature related to the work reported in Chapter 4–6.

3. Manuscript 1: What do people living with chronic pain want from a pain forecast? A research prioritisation study

To understand how a pain forecast would be used by people with chronic pain, and the outcome of a pain forecast that would be useful, I conducted work with people with chronic pain. To understand these factors in a qualitative manner, I conducted a focus group, and to reach a greater number of participants, I conducted a UK-wide survey.

For this paper, I worked closely with coauthor KD, the lead of patient and public involvement at the Centre for Epidemiology Versus Arthritis. I led meetings with all coauthors to plan the conceptualization and methodology of the study. I led the investigation stage of the study (the focus group and survey) and directed two coauthors to support breakout groups in the focus group. I conducted the formal analysis of the data acquired in the study. I created the original draft of the manuscript presented below. All coauthors contributed to the final version of the manuscript. I successfully acquired funding for this study from the Centre for Epidemiology Versus Arthritis fund.

This manuscript has been submitted for peer review to *PLOS ONE* and uploaded to the preprint sever *medRxiv*. The version presented here is the publicly available version on *medRxiv*. Additional files included in the manuscript are provided in Appendix D.

What do people living with chronic pain want from a pain forecast? A research prioritisation study

Authors:

Claire L Little[1], Katie L Druce[1], William G Dixon[1,2], David M Schultz[3,4], Thomas House[5], John McBeth[1,2]

Affiliations:

[1]Centre for Epidemiology Versus Arthritis, University of Manchester, Manchester, United Kingdom

[2] NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

[3]Department of Earth and Environmental Sciences, Centre for Atmospheric Science, University of Manchester, Manchester, United Kingdom

[4] Centre for Crisis Studies and Mitigation, University of Manchester, Manchester, United Kingdom

[5]Department of Mathematics, University of Manchester, Manchester, United Kingdom

Corresponding Author:

Claire Little, claire.little@postgrad.manchester.ac.uk

Abstract:

Background: People with chronic pain report feelings of uncertainty and unpredictability around their future pain. A pain-forecasting model could provide important information to support individuals to manage their daily pain and improve their quality of life. To be useful, the model should be developed with people living with chronic pain. We conducted Patient and Public Involvement (PPI) work, with the aim of this PPI to design the content of a pain-forecasting model by (1) learning participants' priorities in the features of pain provided by a pain forecast and (2) understanding the benefits that participants perceive they would gain from such a forecast.

Methods: A focus group of 12 participants identified potential features, benefits and drawbacks of a pain forecast. In a survey, participants with chronic pain (n = 148) prioritised the identified pain features and perceived benefits.

Results: Focus group participants identified anticipatory anxiety and fears around data-sharing as potential drawbacks. Survey respondents prioritised forecasting of pain flares (68%) and fluctuations in pain severity (64%). Specific priorities about pain flares were the timing of the onset and the severity. Of those surveyed, 75% would use a future pain forecast and 80% perceived making plans (e.g. shopping, social) as a benefit.

Conclusions: For people with chronic pain, the timing of the onset of pain flares, the severity of pain flares and fluctuations in pain severity were prioritised as being key features of a pain forecast, and making plans was prioritised as being a key benefit.

Key Words: Forecast, Chronic Pain, Focus Group, Survey, Research Prioritisation, PPI

Plain English Summary:

Chronic pain is a symptom of many long-term health conditions. People with chronic pain have reported that the severity of their pain is both uncertain and unpredictable. To combat this, we want to build a pain forecast, to predict future pain severity. We hypothesise that a pain forecast would reduce pain-related uncertainty and improve quality of life. It is important that a pain forecast provides useful information to people living with chronic pain. Therefore, this work aimed to understand why participants might use a forecast, and what they would want to see in a pain forecast.

A focus group was conducted to identify features, benefits and drawbacks of a pain forecast. A survey was then conducted to prioritise the features and benefits. Participants of the focus group highlighted concerns around data-sharing and potential anxiety about knowing when pain might happen. Survey participants prioritised a forecast that provided information about pain flares (periods of increased pain severity) and fluctuations in pain severity. The key perceived benefit of a forecast was the ability to make plans (such as shopping and social plans).

Background

Chronic pain (i.e. pain lasting at least three months) is experienced by an estimated 43% of adults in the United Kingdom (1,2). Chronic pain conditions are associated with significant individual and societal burden. They are among the leading causes of disability globally (3). Individuals report that pain interferes with their professional and social lives, affects their relationships, and decreases their quality of life, mood and sleep (4). In the UK, 13.4% of sickness days were due to musculoskeletal conditions in 2021 (5). Although up-to-date figures are scarce, the economic costs of chronic pain are considerable. For example, back pain (a common cause of chronic pain) cost up to £12.3 billion in the UK in 1998 (6) and chronic pain conditions cost 1.5–3% of European GDP in 2012 (7).

The severity of chronic pain is a key driver of outcome, with more severe pain associated with worse outcomes including poorer physical and mental health–related quality of life (8–10), mood (11–13), and social and work participation (14,15). However, the absolute level of pain severity is not the only important driver of outcome. Variability in pain severity is also an important factor. The severity of chronic pain is not stable over time, and individuals experience intra- and inter-daily fluctuations in pain severity and pain flares which are characterised by a rapid increase in pain severity (16–22). People living with chronic pain report that the variability in pain severity is unpredictable, and this unpredictability leads to feelings of uncertainty (23,24) that permeates every sphere of their lives through a decreased ability to work, missed social events, and avoidance in making commitments (25,26). There is a clear desire to reduce the unpredictability of pain severity, with patients often asking how their pain might manifest in the future.

Pain is a complex biopsychosocial phenomenon and predicting variability in pain severity, including pain flares, will be challenging. It will involve identifying and understanding the complex relationship between time-varying biological, psychological and social exposures, discerning how those are associated with changes in pain severity over time, and developing models to forecast those changes. We propose that a personalised pain-forecasting model could reduce pain-related uncertainty by providing predictions of future pain. We have identified factors that are associated with variability in pain severity and could be used as predictors in such a model, including prior pain experience, physical activity (27,28), mood (29,30), sleep quality (31) and environmental exposures (here, the weather) (32).

Recent developments in digital data collection tools offer a solution to capturing these data. Patient-generated health data in chronic pain are already used to track daily symptoms including pain symptoms over time (33), to inform models of care (34) and to facilitate conversations between clinicians and patients (35). Other spheres have shown the feasibility of using patient generated health data to forecast symptoms. For example, individualised prediction models exist for forecasting the diagnosis and prognosis of Covid-19 (36), the presence of anxiety and depression (37), the severity of hayfever symptoms (38) and the level of physical fatigue (39). It is feasible that patient-generated health data could also be used to forecast the variability in pain severity. However, the features that a pain forecasting model should predict are not yet clear.

There are many potential pain features that could be predicted by a pain forecasting model including, for example, the level of forecasted pain severity described as an absolute value, the level of change in forecasted pain severity described as an absolute or proportional increase, the timing of that change, and the variability in pain severity

over time. It is not clear which, if any, of these features people living with chronic pain would prioritise in a pain forecast. Patient and Public Involvement (PPI) is defined as work done *with* members of the public and can be conducted to involve stakeholders in the research process, including in identifying research priorities (40,41). Conducting PPI in the process of developing a pain forecast would ensure that the forecast is suited to the needs and priorities of its users (42). Thus, identifying and prioritising pain features in PPI activities forms the first stage in producing a pain-forecasting model.

The objectives of this PPI work were to design the content of a pain forecast by (1) learning participants' priorities in the features of pain severity provided by a pain forecast and (2) understanding the benefits that participants perceive they would gain from such a forecast.

Methods

Overall study design

Two PPI activities were conducted with individuals with chronic pain. The first PPI activity was a focus group to inform the second PPI activity, a survey of people living with chronic pain. The aim of the focus group was to identify potential pain features that could be produced by a pain forecast and a list of potential benefits of a pain forecast. The aim of the survey was to prioritise these features and benefits in a larger sample of people living with chronic pain. These PPI activities were approved by the Proportionate University Research Ethics Committee at the University of Manchester (Ref: 2021-11862-19751). The activities are reported in line with the GRIPP2 (Guidance for Reporting Involvement of Patients and the Public) checklist (41), which is provided in Additional file 1.

Focus group

A semi-structured focus group was conducted with individuals with chronic pain to produce a list of meaningful pain features that could be provided by a forecast and to understand the perceived potential benefits of a forecast. A focus group was chosen as it allowed us to explore reasons behind the choices and to allow participants to build on each other's ideas (43). A single focus group was conducted due to time and budget constraints.

We sought to recruit up to 12 individuals who were at least 18 years old, who self-reported having a non-cancer chronic-pain condition, lived in the UK, and could read English. Participants were recruited through social media and newsletters of charity organisations related to non-cancer chronic-pain conditions (see Additional file 2) and shared through professional social media accounts of colleagues. Potential participants completed a screening questionnaire, providing demographic information on their gender, ethnic group, age bracket (18–25, 26–45, 46–65 or 66+), self-reported chronic-pain condition(s) from a multiple-choice list and length of time since diagnosis. Participants for the focus group were then selected using purposive sampling, ensuring variation in age, gender, ethnic group, number and type of chronic-pain condition(s) and time since diagnosis. Recruited individuals provided informed written consent and were reimbursed for their time and expenses, in line with PPI guidelines from the National Institute for Health and Care Research (44).

The focus group took place in August 2021 and lasted approximately 90 minutes. Due to the ongoing COVID-19 pandemic, the focus group was held online using Zoom. Three researchers (authors CL, KD and JM) co-facilitated the focus group and made written, anonymised field notes.

The structure of the focus group is provided in Additional file 3. Discussion topics focussed on the pain features of a forecast that participants identified as potentially beneficial, how a forecast could be used in day-to-day life and how the survey (the second PPI activity) should be structured. These discussions led to revisions of the survey, details of which are provided in the results section. The structure of the focus group included a general group-level introduction (facilitated by CL), breakout-room discussions (facilitated by CL, KD and JM), and a final group-level discussion.

Group-level discussions in the focus group were audio-recorded through Zoom, without video recordings, and subsequently transcribed verbatim by CL. Field notes from breakout rooms were made by all facilitators. Participants' views on potential features of a pain forecast, and how a forecast may be used in day-to-day life were noted and subsequently used to inform multiple-choice questions in the survey.

Survey

The second PPI activity was a survey of people living with chronic pain. This survey aimed to learn participants' priorities regarding the potential features and perceived benefits of a pain forecast, using the features and benefits identified by the focus group participants.

The survey was distributed in October and November 2021. The recruitment strategy was identical to that of the focus group. Members of the focus group were not prohibited from completing the survey. Consent was provided electronically at the start of the survey, and only complete surveys were analysed. We aimed to receive at least 100 completed surveys.

The survey collected demographic information and included priority-setting questions and multiple-choice questions. The demographic information collected were participants' gender, age, chronic pain condition(s) and site(s) of pain. Priority-setting questions asked participants to order multiple-choice options in preference order, using a drag-and-drop feature. Multiple-choice questions asked participants to select one or multiple options, with free-text space for further elaboration if required. Priority-setting and multiple-choice questions were written following discussions in the focus group, and so details are deferred to the Results section. The full survey can be viewed in Additional file 4.

Analysis of the survey questions was conducted descriptively. For priority-setting questions, the distribution of respondents prioritising each option as most important to least important was calculated. For multiple-choice questions, the number and percentage of participants selecting each response was calculated. Sensitivity analyses using chi-squared tests were conducted to compare the responses of participants with commonly reported pain conditions to those without.

Due to the small number of free-text responses, they were not directly analysed but are reported following the relevant questions.

Results

Focus group

Demographic data of the 12 participants are provided in Table 1. There were nine females and three males, with all age brackets (18–25, 26–45, 46–65 or 66+) represented. Six participants had been living with chronic pain for at least five years, and five participants had two or more chronic-pain conditions. Chronic-pain conditions

of the participants included osteoarthritis, chronic headache, fibromyalgia, neuropathic pain, rheumatoid arthritis and spondyloarthritis.

Discussions in the focus group were centred on three overarching themes: the pain features, if any, that participants wanted a forecast to provide, the perceived benefits of a forecast and the potential drawbacks of using a forecast in daily life. These themes are discussed in more detail below, followed by the implications they had on the survey questions.

The first overarching theme that participants discussed was the pain features of a forecast. Participants described pain features of interest in relation to periods of pain flares (when pain severity increased for a number of days) and periods of low or no pain severity (when pain severity decreased for a number of days). First, they spoke about the timing of these periods (such as when a period of low pain might begin), as this would have implications on activities that could be undertaken, and may impact how work and social activities could be managed in advance:

> P6 (F66+): "I gave an example of a [colleague] at the moment who's got a frozen shoulder and she'll be coming back on a phased return but the timing of that may be helpful through the [forecast]"

Second, they discussed the duration of pain severity at high or low levels. Knowledge about the duration of pain severity at certain levels would also allow participants to plan activities and relevant interventions. For example, a participant in a breakout room commented that they may visit a chiropractor if their pain was predicted to last for a substantial period of time.

Table 1: Demographics of focus group participants (*n* = 12)	
	Number of participants
Age	
18—25	1
26—45	3
46—65	6
66+	2
Gender	
Female	9
Male	3
Ethnicity	
Asian British	1
Black British	1
British Chinese	1
South Asian	1
White British	7
White British & Black African	1
Number of chronic pain conditions	
1	7
2 or more	5
Chronic pain condition*	
Chronic headache	2
Fibromyalgia	2
Neuropathic pain	2
Osteoarthritis	5
Rheumatoid Arthritis	2
Spondyloarthritis	2
Other	3
Time since diagnosis	
Under 5 years	6
5 or more years	6

*Column exceeds 100% since participants can have multiple chronic pain conditions

Third, participants wanted to know information about their pain-related quality of life and other symptoms (such as fatigue). Participants in breakout rooms noted that these other symptoms were not always correlated to pain severity and wanted to know information about them separately.

> P9 (F46–65): "I think you should maybe focus on quality of life. We all have different levels of pain, different levels of fatigue, we are all different. But what is important for me is totally different to the next person… It's on what you would accept as a quality of life"

The second overarching theme that participants explored were the perceived benefits of a pain forecast. The most commonly reported benefit of a forecast was that it could improve the ability to make plans, including meeting family, planning grocery shopping, planning pharmacological interventions, organising non-pharmacological interventions and adapting work:

> P3 (F26–45): "For me… the biggest advantage would be planning medication… I tend to go for the lowest meds, and then regret it because I'm still in pain and, oh God, now I can't take this, or I could take it but then I have stomach issues, all the rest of it. So… if I could get my drugs more accurate to how it's going to be, my pain medication, my PMR [steroid medication], that would really help, I think."

Another reported benefit was that participants hoped a forecast might support them in understanding triggers of pain, including how variables such as weather, stress and exercise might affect pain severity. In breakout rooms, participants discussed the empowerment granted through understanding their own triggers of pain.

In the third overarching theme, participants identified potential drawbacks of a pain forecast. First, participants highlighted potential mental-health challenges, such as anxiety and stress induced by having information about pain events, including pain flares:

> P3 (F26–45): "There are mental health disadvantages like anticipatory anxiety if the [forecast] tells you you're going to feel rubbish in a week."

Among these concerns, participants voiced fears of a self-fulfilling prophecy if they expected pain severity to increase.

Other participants highlighted mental-health challenges related to inputting pain-severity scores, which may encourage higher focus on the pain that they are trying to manage.

Second, participants were anxious about the potential implications of data collection during a pain forecast and fears of data sharing with employers and government officials:

> P7 (F46–65): "Who has access to the data? I think it would put a lot of people off if people thought that employers are going to have access to this data."

> P6 (F66+): Would it be "used by occupational health departments in organisations?"

Based on the discussion of the focus group, priority-setting and multiple-choice questions for the survey were written. Of the three overarching themes, questions were developed regarding the potential features and benefits of a pain forecast. As the drawbacks related to the implementation of a pain forecast, these were not included in the present survey. All suggested pain features of the focus group discussion were

included, asking survey participants to prioritise the importance of timing, duration and symptoms during periods of low pain and pain flares. All reported benefits were included, asking survey participants to select which benefits applied to them regarding planning, applying pharmacological and non-pharmacological interventions and understanding triggers of pain.

Survey

There were 148 respondents to the survey. Demographic information and data regarding chronic pain condition can be found in Table 2. Approximately 90% of respondents were female and over three quarters were aged between 36 and 65. The most commonly reported chronic-pain conditions were fibromyalgia (46%) and osteoarthritis (33%). Sensitivity analysis discovered no differences between the responses of participants with these common pain conditions and other pain conditions (see Additional file 5). Nine participants reported only 'other' pain conditions. These participants reported ankylosing spondylitis, psoriatic arthritis, scleroderma, systemic lupus erythematosus and juvenile arthritis.

Results of the multiple-choice question *"Which of the following would you like a pain forecast to provide for you?"* are shown in Table 3. The most commonly selected features were pain flares (68%) and fluctuations in pain severity (64%). Features of pain severity on an ordinal scale (47%) and periods of low pain (35%) were less commonly selected.

Table 2: Demographics and chronic pain condition of survey respondents (*n* = 148)			
		Number of participants	Percentage of participants
Age	18–25	11	7.4%
	26–35	19	12.8%
	36–45	34	23.0%
	46–55	27	18.2%
	56–65	40	27.0%
	66+	17	11.5%
Gender	Male	13	8.8%
	Female	134	90.5%
	Non-binary/third gender	1	0.7%
Chronic pain condition*	Fibromyalgia	68	46.0%
	Osteoarthritis	49	33.1%
	Chronic headache	32	21.6%
	Neuropathic pain	28	18.9%
	Rheumatoid Arthritis	19	12.8%
	Spondyloarthritis	17	11.5%
	Unspecific Arthritis	12	8.1%
	Other	69	46.6%
*Column exceeds 100% since participants can have multiple chronic pain conditions			

Table 3: "Which of the following would you like a pain forecast to provide for you?"		
Pre-specified response	**Number of participants who selected response**	**Percentage of participants who selected response**
Information about a pain flare	100	67.6%
Information about fluctuations in pain severity	94	63.5%
Information about pain severity on a scale of 1 to 5	70	47.3%
Information about a period of low/no pain severity	51	34.5%
Other/None	13	8.8%
Responses to the question: "Which of the following would you like a pain forecast to provide for you?" Participants could select more than one option.		

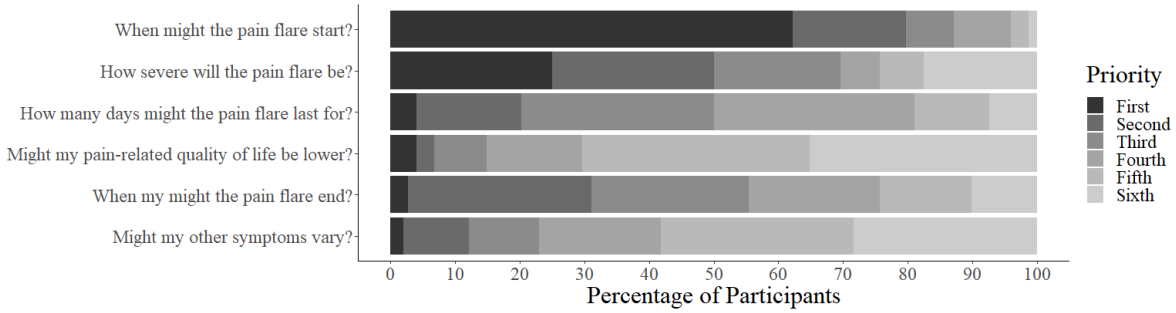
Only two relevant free-text responses were provided to this question, both referring to fluctuations in pain severity:

> "Will my future pain graphs differ from those in the past?"

> "Compare it to half hour ago, a couple of hours ago, yesterday etc with a higher lower method"

Results of the priority-setting question "*If we could predict a pain flare, what specific information would you want to know?*" are shown in Figure 1. The respondents ranked six statements. Onset of a pain flare was the first priority for 92 (62%) of respondents. Severity of a pain flare was the first or second priority for 50% of respondents. Pain-related quality of life and variation in other symptoms were given fifth or sixth priority by 70% and 58% of respondents, respectively. Responses to this question highlight that the onset of a pain flare and severity of a pain flare are clear priorities for respondents.

Figure 1: Survey respondents' priorities of the pre-specified responses relating to pain flares.



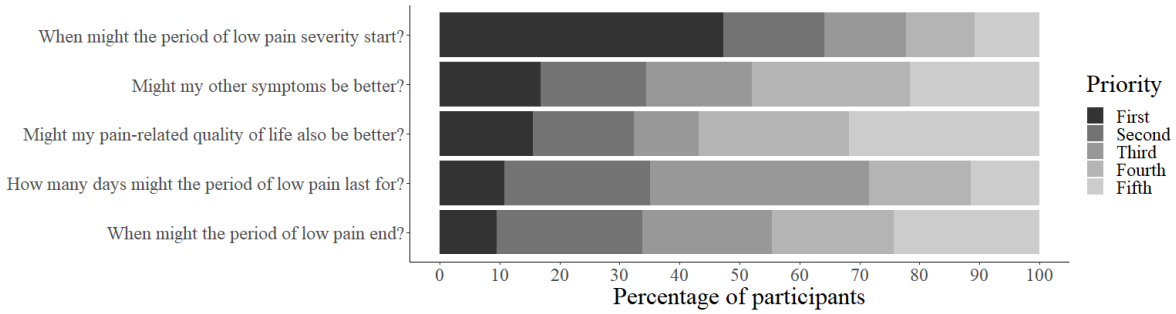
Respondents were prompted with the question: "If we could predict a pain flare, what specific information would you want to know?". Percentages of participants ranking each statement as their first, second, third, fourth, fifth or sixth priority are reported.

In free-text responses, nine respondents highlighted that they also wanted information about the triggers of their pain flare. Specific triggers that were cited were hormonal cycles, weather, environment and mood. One participant wanted information about the acceleration of the pain flare and one wanted information about medication to take during a flare.

Results of the priority-setting question "*If we could predict a period of low pain severity, what specific information would you want to know?*" are shown in Figure 2. The respondents ranked five statements. Onset of a period of low pain severity was first priority for 70 (47%) respondents. Respondents did not show great variability among the other responses. First or second priority was given to the duration of the period by 36% of participants, to variation in other symptoms by 35% of participants, to the end of the period by 34% of participants, and to pain-related quality of life by 32% of

participants. There was therefore only a clear priority for information about the onset of the period of low pain severity.

Figure 2: Survey respondents' priorities of the pre-specified responses relating to periods of low pain severity.



Respondents were prompted with the question: " If we could predict a period of low pain severity, what specific information would you want to know?". Percentages of participants ranking each statement as their first, second, third, fourth or fifth priority are reported.

Of the free text responses, eight respondents referred to wanting information about the triggers of their low pain (e.g. barometric pressure) or variables that they could control (e.g. exercise). One participant wanted to understand how typical their experience is among others, one wanted information about treatments, and one wanted information about specific days.

Participants were asked the multiple-choice question: "*If a pain forecast could provide useful information for you, do you think that you would use a pain forecast?"* (Table 4). Of those surveyed, 75% would use a pain forecast.

Table 4: "Do you think that you would use a pain forecast?"		
Response	**Number of participants who selected response**	**Percentage of participants who selected response**
Definitely not	0	0%
Probably not	11	7.4%
Might or might not	24	16.2%
Probably yes	63	42.6%
Definitely yes	50	33.8%
Responses to the question: "If a pain forecast could provide useful information for you, do you think that you would use a pain forecast?"		

All participants were also asked: "*What would you use a pain forecast for?"* (Table 5). The most common reasons were making plans (83%) and understanding individual triggers of chronic pain (76%). In addition, 47% and 31% of respondents would use a pain forecast to plan pharmacological and non-pharmacological interventions respectively. Therefore, making plans and understanding triggers are highlighted as the most likely benefits, although planning pharmacological and non-pharmacological interventions may also be of interest to a number of users.

Table 5: "What would you use a pain forecast for?"		
Pre-specified response	**Number of participants who selected response**	**Percentage of participants who selected response**
To help make plans (e.g. shopping, social)	123	83.1%
To understand the triggers of my pain	113	76.4%
To know when my pain severity might be better/worse	92	62.2%
To help plan non-pharmacological interventions	70	47.3%
To help choose which medication to take	46	31.1%
Responses to the question "What would you use a pain forecast for?" Participants could select more than one option.		

Relevant free-text responses were:

"To improve my overall self management of my conditions"

"To let work know times where I might need time off to recover so that it's not out of the blue for them and they can prepare for me to be off if I need to"

"To analyse the development of my condition"

"Exercise planning"

"To help me understand my condition more"

"To look forward to some good times!"

Discussion

In our PPI activities, the content of a pain forecasting was discussed with people with chronic pain in a focus group and subsequently prioritised in a UK-based survey. We report that survey participants highlighted an interest in a forecast involving periods of pain flares and fluctuations in pain severity. Within a priority-setting question about pain flares, the most highly ranked features were the timing of the start of their flare and the severity of their flare. Taken together, the participants in these PPI activities were clear in their interest in receiving information about the timing of the start of their flare, the severity of their flare, and fluctuations in their pain severity.

Over three quarters of the surveyed population reported that a forecast would help them to make plans and assist in understanding triggers of their pain. Respondents would also use a pain forecast to understand when their pain may be better or worse, to guide the use of non-pharmacological and pharmacological interventions. However, the focus group also highlighted concerns about the use of a forecast in day-to-day life, including anticipatory anxiety and sharing of personal data.

Developing the survey following the focus group ensured that the themes presented were identified as being important and meaningful to individuals with chronic pain. The survey results can also be interpreted in the context of topics raised in the focus group, including how a forecast might be used to support going back to work or to plan pharmacological interventions. As a result, the PPI activities allow a pain forecast to be developed with the feedback from stakeholders.

There are limitations to the PPI activities that should be considered. The representation of different conditions in our survey may have been impacted by the charities that

shared the advertisements with their members, perhaps explaining the high prevalence of fibromyalgia among our respondents. This would impact the results if participants with certain conditions had different priorities to other people with chronic pain, and those conditions were over-represented in the surveyed population. However, sensitivity analyses among the subset of participants with fibromyalgia compared to the subset of participants without fibromyalgia, and the subset of participants with osteoarthritis compared to the subset of participants without osteoarthritis found no differences in the reported responses.

Recruitment advertisements clarified that participants would be commenting on a pain forecast, and respondents therefore had an interest in commenting on this topic. A large proportion of our survey participants were interested in a pain forecast and so the results may be less generalisable to those people who would initially be unsure or less inclined to use a forecast.

Both PPI activities recruited participants online, primarily due to the ongoing COVID-19 pandemic. A pain forecast may be implemented in a future digital intervention and users would then be required to access the internet. However, our work did not include participants who may be less inclined to use the internet for reasons including access and digital literacy. Our findings are therefore not generalisable to these populations and if a digital intervention is developed, work with these populations should be considered.

Previous work has reported on the longer-term prognosis of chronic pain. For example, some studies have followed people with chronic pain over several months or years and identified different trajectories of pain severity among people with chronic pain (45,46). Other studies have identified prognostic factors associated with chronic pain outcomes

(47). However, the participants in our focus group and survey have highlighted the importance of forecasting pain on a shorter-term, to support daily activities.

The benefits of a pain forecast extend previous work. Flurey et al. (23) reported that patients expressed frustration at the unpredictability of pain flares and this led to participants cancelling or altering plans. Fullen et al. (26) also found that individuals with chronic low-back pain reported missing out on social events and avoided making commitments, due to the unpredictability of their pain. Our work highlighted that respondents would value a forecast that reduced the unpredictability of their pain, particularly around the timing and severity of pain flares. Our participants highlighted the importance of making plans as a key benefit of a forecast, likely due to the frustration previously reported which has resulted in avoidance of making plans.

The drawbacks highlighted are also consistent with previous work. Among the challenges in collection and analysis of patient generated health data, privacy concerns have previously been highlighted (48), in line with concerns of focus group participants. Any future mobile application of a pain forecast should follow standards of privacy and security (49) and clearly communicate these to users. Furthermore, as pain is widely accepted within the biopsychosocial model (50), and rumination and catastrophizing are associated with increased pain severity (51), concerns around anticipatory anxiety should be thoughtfully considered.

Conclusions

These PPI activities indicated a high level of interest in a pain forecast and our participants were clear that pain features should include the timing of the start of a pain flare, the severity of a pain flare, and fluctuations in pain severity. Future work will

develop a statistical pain forecasting model, to predict these identified features. As one of the key benefits of a pain forecast is the identification of triggers, a future model should be interpretable by its users. Drawbacks highlighted in the focus group, such as the impact of anticipatory anxiety should also be considered during the production of a forecast. Wider interest will be determined in the future, based on the uptake of a forecast and continued involvement of stakeholders and evaluation of a forecast will ensure that priorities indicated by participants translate into real value.

List of Abbreviations

PPI: Patient and Public Involvement

Declarations

Ethics approval and consent to participate: This study was approved by the Proportionate University Research Ethics Committee at the University of Manchester (Ref: 2021-11862-19751). Informed written consent was provided by members of the focus group. Informed electronic consent was provided by respondents to the survey. This procedure to take informed consent electronically from the respondents was approved by the Proportionate University Research Ethics Committee at the University of Manchester. All methods were carried out in accordance with relevant guidelines and regulations (Declaration of Helsinki).

Availability of data and materials: The datasets generated and analysed during the current study are not publicly available as data were collected for patient and public involvement activities and consent was obtained for the sharing of anonymous quotes and aggregated data only. However, data are available from the corresponding author on reasonable request.

Competing interests: WGD has received consultancy fees from Google, and DMS has received consultancy fees from Palta, both unrelated to this work. All other authors have no conflicts of interest to declare.

Funding: This work was supported by infrastructure support from the Centre for Epidemiology Versus Arthritis (grant number 21755).

TH receives funding from the Royal Society (grant number INF/R2/180067) and the Alan Turing Institute for Data Science and Artificial Intelligence.

Authors' contributions: All authors contributed to the conception of the study. CL, KLD and JMcB designed the study and conducted the focus group and survey. CL analysed the results and wrote the first draft of the article. All authors contributed to revising the final article.

Acknowledgements: We are grateful for the contributions of the members of the Patient and Public Involvement focus group and the respondents to the survey.

References

1. McQueenie R, Jani BD, Siebert S, McLoone P, McCowan C, Macdonald S, et al. Prevalence of chronic pain in LTCs and multimorbidity: A cross-sectional study using UK Biobank. J Multimorb Comorbidity. 2021 Jan 1;11:263355652110058.

2. Fayaz A, Croft P, Langford RM, Donaldson LJ, Jones GT. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. BMJ Open. 2016 May;6(6):e010364.

3. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet. 2015 Aug;386(9995):743–800.

4. Hadi MA, McHugh GA, Closs SJ. Impact of Chronic Pain on Patients' Quality of Life: A Comparative Mixed-Methods Study. J Patient Exp. 2019 Jun;6(2):133–41.

5. Office for National Statistics. Sickness absence in the UK labour market: 2021 [Internet]. 2022 Apr [cited 2022 Aug 9]. Available from: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles/sicknessabsenceinthelabourmarket/2021

6. Maniadakis N, Gray A. The economic burden of back pain in the UK. Pain. 2000 Jan;84(1):95–103.

7. Barham L. Economic Burden of Chronic Pain Across Europe. J Pain Palliat Care Pharmacother. 2012 Mar 7;26(1):70–2.

8. Leadley RM, Armstrong N, Reid KJ, Allen A, Misso KV, Kleijnen J. Healthy aging in relation to chronic pain and quality of life in Europe. Pain Pract Off J World Inst Pain. 2014 Jul;14(6):547–58.

9. Mutubuki EN, Beljon Y, Maas ET, Huygen FJPM, Ostelo RWJG, van Tulder MW, et al. The longitudinal relationships between pain severity and disability versus health-related quality of life and costs among chronic low back pain patients. Qual Life Res Int J Qual Life Asp Treat Care Rehabil. 2020 Jan;29(1):275–87.

10. Jensen MP, Chodroff MJ, Dworkin RH. The impact of neuropathic pain on health-related quality of life: review and implications. Neurology. 2007 Apr 10;68(15):1178–82.

11. Fonseca-Rodrigues D, Rodrigues A, Martins T, Pinto J, Amorim D, Almeida A, et al. Correlation between pain severity and levels of anxiety and depression in osteoarthritis patients: a systematic review and meta-analysis. Rheumatol Oxf Engl. 2021 Dec 24;61(1):53–75.

12. Bonilla-Jaime H, Sánchez-Salcedo JA, Estevez-Cabrera MM, Molina-Jiménez T, Cortes-Altamirano JL, Alfaro-Rodríguez A. Depression and Pain: Use of Antidepressants. Curr Neuropharmacol. 2022;20(2):384–402.

13. Stubbs B, Vancampfort D, Veronese N, Thompson T, Fornaro M, Schofield P, et al. Depression and pain: primary data and meta-analysis among 237 952 people across 47 low- and middle-income countries. Psychol Med. 2017 Dec;47(16):2906–17.

14. Froud R, Patterson S, Eldridge S, Seale C, Pincus T, Rajendran D, et al. A systematic review and meta-synthesis of the impact of low back pain on people's lives. BMC Musculoskelet Disord. 2014 Feb 21;15:50.

15. Cancelliere C, Donovan J, Stochkendahl MJ, Biscardi M, Ammendolia C, Myburgh C, et al. Factors affecting return to work after injury or illness: best evidence synthesis of systematic reviews. Chiropr Man Ther. 2016;24(1):32.

16. Austin L, Sharp CA, van der Veer SN, Machin M, Humphreys J, Mellor P, et al. Providing 'the bigger picture': benefits and feasibility of integrating remote monitoring from smartphones into the electronic health record. Rheumatology. 2020 Feb 1;59(2):367–78.

17. Allen KD, Coffman CJ, Golightly YM, Stechuchak KM, Keefe FJ. Daily pain variations among patients with hand, hip, and knee osteoarthritis. Osteoarthritis Cartilage. 2009 Oct;17(10):1275–82.

18. Bartley EJ, Robinson ME, Staud R. Pain and Fatigue Variability Patterns Distinguish Subgroups of Fibromyalgia Patients. J Pain. 2018 Apr;19(4):372–81.

19. Toussaint L, Vincent A, McAllister SJ, Whipple M. Intra- and Inter-Patient Symptom Variability in Fibromyalgia: Results of a 90-Day Assessment: Symptom Variability in Fibromyalgia. Musculoskeletal Care. 2015 Jun;13(2):93–100.

20. Harris RE, Williams DA, McLean SA, Sen A, Hufford M, Gendreau RM, et al. Characterization and consequences of pain variability in individuals with fibromyalgia. Arthritis Rheum. 2005 Nov;52(11):3670–4.

21. Vives-Mestres M, Casanova A, Hershey AD, Orr SL. Perceived stress and pain severity in individuals with chronic migraine: A longitudinal cohort study using daily prospective diary data. Headache J Head Face Pain. 2021 Sep;61(8):1245–54.

22. Houle TT, Penzien DB, Rains JC. Time-Series Features of Headache: Individual Distributions, Patterns, and Predictability of Pain. Headache J Head Face Pain. 2005 May;45(5):445–58.

23. Flurey CA, Morris M, Richards P, Hughes R, Hewlett S. It's like a juggling act: rheumatoid arthritis patient perspectives on daily life and flare while on current treatment regimes. Rheumatology. 2014 Apr 1;53(4):696–703.

24. Darlow B, Brown M, Thompson B, Hudson B, Grainger R, McKinlay E, et al. Living with osteoarthritis is a balancing act: an exploration of patients' beliefs about knee pain. BMC Rheumatol. 2018 Dec;2(1):15.

25. Hutchings A, Calloway M, Choy E, Hooper M, Hunter DJ, Jordan JM, et al. The Longitudinal Examination of Arthritis Pain (LEAP) study: relationships between weekly fluctuations in patient-rated joint pain and other health outcomes. J Rheumatol. 2007 Nov;34(11):2291–300.

26. Fullen B, Morlion B, Linton SJ, Roomes D, Griensven J, Abraham L, et al. Management of chronic low back pain and the impact on patients' personal and professional lives: Results from an international patient survey. Pain Pract. 2022 Feb 25;papr.13103.

27. Kolber MR, Ton J, Thomas B, Kirkwood J, Moe S, Dugré N, et al. PEER systematic review of randomized controlled trials: Management of chronic low back pain in primary care. Can Fam Physician Med Fam Can. 2021 Jan;67(1):e20–30.

28. Antcliff D, Campbell M, Woby S, Keeley P. Activity Pacing is Associated With Better and Worse Symptoms for Patients With Long-term Conditions. Clin J Pain. 2017 Mar;33(3):205–14.

29. George SZ, Parr JJ, Wallace MR, Wu SS, Borsa PA, Dai Y, et al. Biopsychosocial Influence on Exercise-Induced Injury: Genetic and Psychological Combinations Are Predictive of Shoulder Pain Phenotypes. J Pain. 2014 Jan;15(1):68–80.

30. Read JR, Sharpe L, Modini M, Dear BF. Multimorbidity and depression: A systematic review and meta-analysis. J Affect Disord. 2017 Oct 15;221:36–46.

31. McBeth J, Dixon WG, Moore SM, Hellman B, James B, Kyle SD, et al. Sleep Disturbance and Quality of Life in Rheumatoid Arthritis: Prospective mHealth Study. J Med Internet Res. 2022 Apr 22;24(4):e32825.

32. Yimer BB, Schultz DM, Beukenhorst AL, Lunt M, Pisaniello HL, House T, et al. Heterogeneity in the association between weather and pain severity among patients with chronic pain: a Bayesian multilevel regression analysis. PAIN Rep. 2022 Jan;7(1):e963.

33. Druce KL, Dixon WG, McBeth J. Maximizing Engagement in Mobile Health Studies: Lessons Learned and Future Directions. Rheum Dis Clin North Am. 2019 May;45(2):159–72.

34. Gandrup J, Selby DA, van der Veer SN, Mcbeth J, Dixon WG. Using patient-reported data from a smartphone app to capture and characterize real-time patient-reported flares in rheumatoid arthritis. Rheumatol Adv Pract. 2022;6(1):rkac021.

35. Lordon RJ, Mikles SP, Kneale L, Evans HL, Munson SA, Backonja U, et al. How patient-generated health data and patient-reported outcomes affect patient-clinician relationships: A systematic review. Health Informatics J. 2020 Dec;26(4):2689–706.

36. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ. 2020 Apr 7;369:m1328.

37. Moshe I, Terhorst Y, Opoku Asare K, Sander LB, Ferreira D, Baumeister H, et al. Predicting Symptoms of Depression and Anxiety Using Smartphone and Wearable Data. Front Psychiatry. 2021 Jan 28;12:625247.

38. Voukantsis D, Berger U, Tzima F, Karatzas K, Jaeger S, Bergmann KC. Personalized symptoms forecasting for pollen-induced allergic rhinitis sufferers. Int J Biometeorol. 2015 Jul;59(7):889–97.

39. Hajifar S, Sun H, Megahed FM, Jones-Farmer LA, Rashedi E, Cavuoto LA. A forecasting framework for predicting perceived fatigue: Using time series methods to forecast ratings of perceived exertion with features from wearable sensors. Appl Ergon. 2021 Jan;90:103262.

40. INVOLVE. Briefing notes for researchers: involving the public in NHS, public health and social care research. INVOLVE, Eastleigh; 2012.

41. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. BMJ. 2017 Aug 2;j3453.

42. O'Cathain A, Croot L, Sworn K, Duncan E, Rousseau N, Turner K, et al. Taxonomy of approaches to developing interventions to improve health: a systematic methods overview. Pilot Feasibility Stud. 2019 Dec;5(1):41.

43. Kitzinger J. Qualitative Research: Introducing focus groups. BMJ. 1995 Jul 29;311(7000):299–302.

44. National Institute for Health and Care Research. Reward and recognition for public contributors - a guide to the payment of fees and expenses [Internet]. 2021 [cited 2021 Jul 1]. Available from: https://www.nihr.ac.uk/documents/reward-and-recognition-for-public-contributors-a-guide-to-the-payment-of-fees-and-expenses/12248

45. Kongsted A, Kent P, Axen I, Downie AS, Dunn KM. What have we learned from ten years of trajectory research in low back pain? BMC Musculoskelet Disord. 2016 21;17:220.

46. Previtali D, Andriolo L, Di Laura Frattura G, Boffa A, Candrian C, Zaffagnini S, et al. Pain Trajectories in Knee Osteoarthritis-A Systematic Review and Best Evidence Synthesis on Pain Predictors. J Clin Med. 2020 Sep 1;9(9):E2828.

47. Nieminen LK, Pyysalo LM, Kankaanpää MJ. Prognostic factors for pain chronicity in low back pain: a systematic review. PAIN Rep. 2021;6(1):e919.

48. Bourke A, Dixon WG, Roddam A, Lin KJ, Hall GC, Curtis JR, et al. Incorporating patient generated health data into pharmacoepidemiological research. Pharmacoepidemiol Drug Saf. 2020 Dec;29(12):1540–9.

49. Mandl KD, Mandel JC, Kohane IS. Driving Innovation in Health Systems through an Apps-Based Information Economy. Cell Syst. 2015 Jul;1(1):8–13.

50. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). Pain. 2019 Jan;160(1):19–27.

51. Meints SM, Edwards RR. Evaluating psychosocial contributions to chronic pain outcomes. Prog Neuropsychopharmacol Biol Psychiatry. 2018 Dec;87:168–82.

4. Manuscript 2: Clustering trajectories of pain severity using daily data from a mobile health study

As discussed in Section 1.3, pain severity follows heterogeneous patterns, and clusters can provide understanding of some of the common trajectories. Further, understanding the frequency of movement between these clusters would highlight the feasibility of creating a forecast that looked at the movement between these clusters. This chapter reports on work conducted to identify common pain clusters and the transitions between them.

I led meetings with my coauthors to conceptualize this study and agree on the methodological approach. The data were curated in previous work by WGD and his team. I implemented the methods and agreed algorithms using computer code in *R*. I conducted the formal analysis. I then wrote the original draft of the manuscript. All coauthors reviewed and edited the manuscript and agreed to its submission for peer review.

The following manuscript is currently under peer-review at *JMIR*. Supplementary materials submitted with this manuscript are provided in Appendix E.

Clustering trajectories of pain severity using daily data from a mobile health study

Paper Type: Original Paper

Authors:
Claire L Little, MA
ORCID: 0000-0003-2602-0491
Centre for Epidemiology Versus Arthritis,
University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

Prof David M Schultz, PhD
ORCID: 0000-0003-1558-6975
Department of Earth and Environmental Sciences,
Centre for Atmospheric Science,
University of Manchester, Manchester, United Kingdom
Centre for Crisis Studies and Mitigation
University of Manchester, Manchester, United Kingdom

Prof Thomas House, PhD
ORCID: 0000-0001-5835-8062
Department of Mathematics
University of Manchester, Manchester, United Kingdom

Prof William G Dixon, MRCP, PhD
ORCID: 0000-0001-5881-4857
Centre for Epidemiology Versus Arthritis,
University of Manchester, Manchester, United Kingdom
NIHR Manchester Musculoskeletal Biomedical Research Unit,
Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Prof John McBeth, PhD
ORCID: 0000-0001-7047-2183
Centre for Epidemiology Versus Arthritis,
University of Manchester, Manchester, United Kingdom
NIHR Manchester Musculoskeletal Biomedical Research Unit,
Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Corresponding Author: Claire L Little, claire.little@postgrad.manchester.ac.uk

Abstract

Background: People with chronic pain have highlighted a need to forecast variability in their pain severity. We propose a forecasting model for both short-term variability (e.g. daily fluctuations) and longer-term variability (e.g. weekly patterns). For development of this model, clusters of weekly trajectories of pain severity are required, so that future work can predict between-cluster movement and within-cluster variability in pain severity. This paper aims to understand clusters of common weekly patterns as a first stage in the development of a pain-forecasting model.

Methods: Data from a population-based mobile health (mHealth) study were used to compile weekly pain trajectories (n = 21,919) and clustered using a k-medoids algorithm. Sensitivity analyses tested the impact of assumptions related to the ordinal and longitudinal structure of the data. The characteristics of people within clusters were examined and a transition analysis was conducted to understand the movement of people between consecutive weekly clusters.

Results: Four clusters of weekly pain severity were identified representing trajectories of no or low pain (n = 1714), mild pain (n = 8246), moderate pain (n = 8376), and severe pain (n = 3583). Sensitivity analyses indicated a four-cluster solution remained suitable under changing assumptions, and resulting clusters were similar to the main analysis, with at least 85% of trajectories belonging to the same cluster as the main analysis. Men spent longer (7.9% of weeks) in the "no or low pain" cluster than women (6.5% of weeks). Younger people (17–24 year olds) spent longer (28.3% of weeks) in the "severe pain" cluster than those aged 65–86 years (9.8% of weeks). People with fibromyalgia (31.5% of weeks) and neuropathic pain (31.1% of weeks) spent longer in the "severe pain" cluster than other conditions, and people with rheumatoid arthritis

spent longer (7.8% of weeks) in the “no or low pain” cluster than other conditions. There were 12,267 pairs of consecutive weeks which contributed to the transition analysis. The empirical percentage remaining in the same cluster across consecutive weeks was 66%. When movement between clusters occurred, the highest percentage of movement was to an adjacent cluster.

Conclusions: The clusters of pain severity identified in this study provide a parsimonious description of the weekly experiences of people with chronic pain. These clusters could be used for future study of between-cluster movement and within-cluster variability, in order to develop accurate and stakeholder-informed pain forecasting tools.

Keywords: mHealth; Pain; Cluster; Trajectory; *k*-medoids; Transition; Forecast

Introduction

Chronic pain (i.e. pain lasting three or more months) is a common symptom of many long-term health conditions [1, 2] and is associated with poor quality of life, poor health outcomes, and low participation in work and social activities [3, 4]. In a meta-synthesis of qualitative studies, patients reported that unpredictability was one of the main features of their pain experience [5]. This unpredictability is likely due to the large amount of day-to-day variability that individuals with chronic pain experience [6, 7], and results in feelings of uncertainty about future pain and a lack of control over symptoms [5, 8]. People living with chronic pain have reported that a pain forecast would reduce unpredictability and could be used to support planning daily activities, such as shopping, chores and social participation [9, 10]. In a research prioritisation study, 75% of respondents to a survey said they would use a pain forecast, and prioritised a model that predicted daily fluctuations (i.e. relatively short-term variability) and pain flares (patterns across multiple days).

We propose three stages to develop a pain forecast (Figure 1). Stage one identifies common weekly trajectories of pain severity using cluster analyses. Stage two investigates day-to-day variability in trajectories of pain severity for individuals within each cluster. These first two stages provide a better understanding of an individual's pain experiences. Stage 3 predicts for an individual their movement between clusters of pain severity across consecutive weeks and future within-cluster day-to-day variability. This study focusses on the first of these stages: clustering trajectories of pain severity.

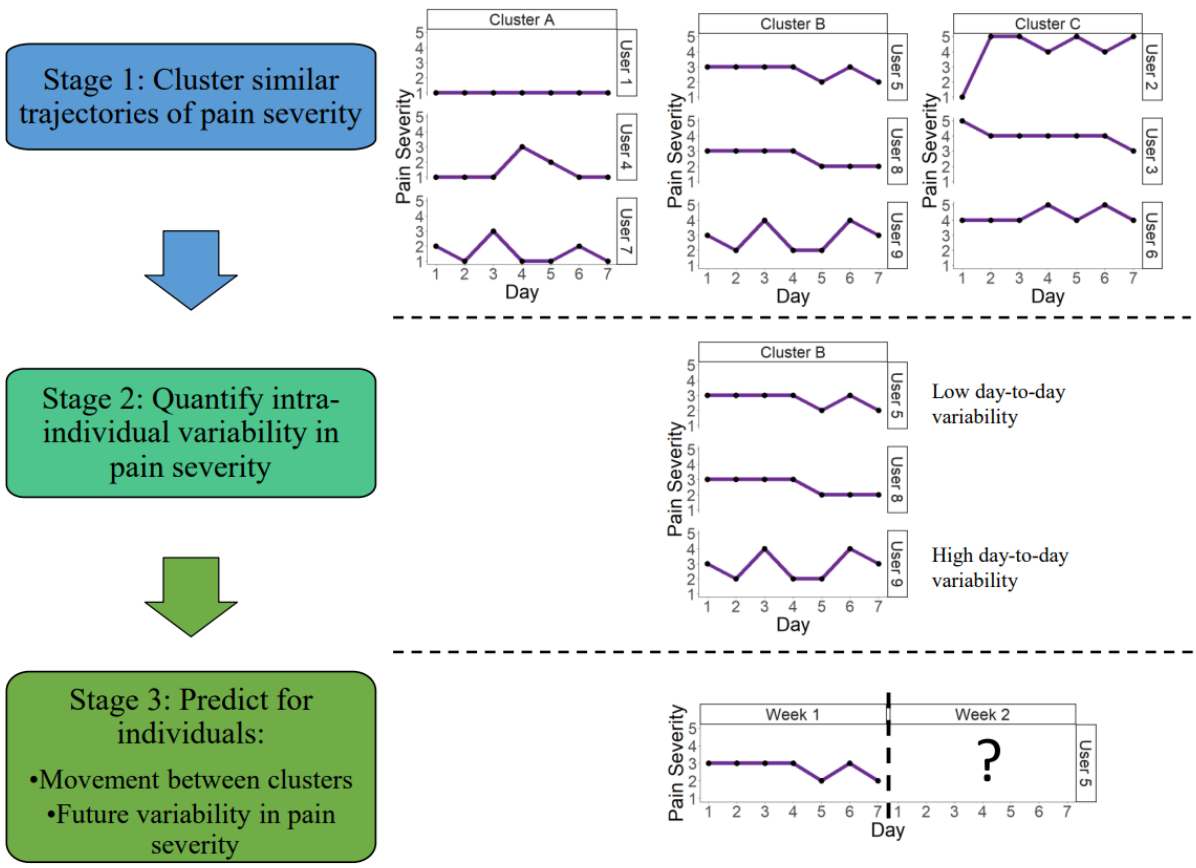


Figure 1: Three stages to build a pain forecast using data from an mHealth study. Data used in this figure are for illustrative purposes only, to provide one example of how data may be used in the pipeline of developing a pain forecast. First, data are clustered to identify common trajectories of weekly pain severity. Second, the remaining variability is explored for each trajectory within a cluster. The process is repeated for each cluster. Third, movement between clusters on consecutive weeks and the amount of day-to-day variability are predicted for an individual. The process is repeated for each individual.

Previous studies have identified patterns of variability in pain severity by clustering pain trajectories among individuals with chronic pain. These studies have often used sparse data on pain severity collected once per week (e.g. [11]), once per month (e.g. [12]) or less frequently (e.g. [13]). These studies inform our understanding of longer-term experiences of chronic pain, but not the day-to-day variability in pain severity that is important to patients. Recent advances

in mobile health (mHealth) methods that support the collection of data in the people's own environments [14, 15], often using their own devices such as smartphones and tablets [16], offer the opportunity to capture daily pain severity data.

Once collected, there are several challenges to overcome in clustering daily pain severity data for use in a pain-forecasting model. First, patient-generated health data are often collected on an ordinal scale. However, equal intervals between responses cannot be assumed, and using metric models to analyse ordinal data can lead to errors [17]. Second, data collected are longitudinal, and algorithms used for clustering should respect this longitudinal feature of the data. Third, clusters of pain severity that will be used in a pain-forecasting model should be interpretable to end-users. To address these challenges requires that we identify and use a suitable method for clustering patient-generated health data. Any assumptions made about the data should be tested in sensitivity analyses to ensure robustness. Observing substantial movement between clusters would suggest the feasibility of forecasting cluster movement in future work. Therefore, understanding the characteristics of individuals who contribute to different clusters and how those individuals move between clusters over time will aid end-user interpretability.

The aim of this study was to understand pain severity clusters in people living with chronic pain. The specific objectives were to (1) use a suitable algorithm to identify the optimum number of clusters of pain trajectories, (2) conduct sensitivity analyses to test assumptions made when clustering data, (3) examine the characteristics of people within clusters, and (4) describe the movement of people between different clusters over time.

Methods

Data source

This study is a secondary analysis of a population-based mHealth study, Cloudy with a Chance of Pain [14, 18, 19]. Study participants were recruited between January 2016 and January 2017. Data collection ended in April 2017. Inclusion criteria required participants to have chronic pain, be aged ≥ 17 years, live in the UK and own an Android or iOS smartphone. Participants

downloaded a mobile phone application (app) and provided demographic information including their sex, year of birth and pain condition(s). Daily reports of ten variables were collected, including pain severity. Participants were asked: "How severe was your pain today?" and responded by selecting "no pain" (score 1), "mild pain" (2), "moderate pain" (3), "severe pain" (4) or "very severe pain" (5). Daily data could be contributed for a maximum of 15 months, with participants requested to track symptoms for six months. In total, 10,483 people downloaded the app and recorded their demographic information and at least one record of pain severity. These participants were 83% female, with a mean age of 51 years (standard deviation (s.d.) 12.5 years).

Data preparation

For this study, weekly trajectories of pain-severity data were used. Trajectories beginning on a Monday were identified, to align data across multiple respondents. This alignment introduced a structure to the data based on the work week, to mitigate the impact of individuals entering the study at different times, and to deal with day-of-the-week effects. A complete participant-week was defined as complete pain-severity data contributed by a single participant during a single calendar week (Monday–Sunday) (Figure 2). Pain-severity data from a complete participant-week were included in the analysis if (a) the participant had joined the study on, or before, the Monday, (b) the participant had remained in the study on, or after, the following Sunday and (c) the participant had provided complete pain-severity data (i.e. one pain severity score on each of the seven days). Multiple complete participant-weeks could be included in the analysis for each participant (up to 64 weeks due to the length of the study).

Statistical methods

Identifying the optimal number of clusters

Previous studies have used a range of methods to cluster pain severity, including *k*-means clustering [20, 21], hierarchical cluster analysis [22], growth mixture modelling [23–26], latent class growth analysis [e.g. 10–12, 26, 27] and group-based trajectory analysis [29–34]. None of

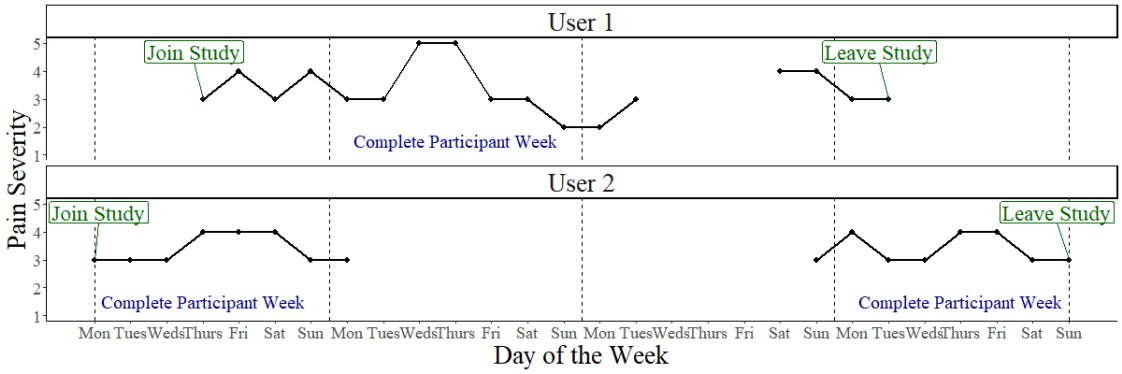


Figure 2: Example selections of complete trajectory-weeks for two participants. The participants join and leave the study at different times. One complete participant-week from User 1 is included in the analysis. Two complete participant-weeks from User 2 are included in the analysis.

these methods can be used to respect both the ordinal and longitudinal nature of the outcome variable. Moreover, clustering of ordinal and longitudinal data in a computationally inexpensive way remains a major unsolved methodological challenge. In this study, we relax the longitudinal assumption of the data by ignoring the ordering of the data (e.g., that Monday directly precedes Tuesday in each trajectory).

To identify the optimal number of clusters, data were summarised in feature vectors, compared using the Manhattan (ℓ_1) distance measure, and clustered using an adaptation of the *k*-medoids algorithm, detailed below. The feature vectors were 7-dimensional with entries representing the pain-severity data on each of the seven days in a complete participant week. Using the data directly in this way ensured that feature vectors remained interpretable. The differences between feature vectors were found by calculating the Manhattan distance through entry-wise summation of absolute differences. The Manhattan distance is a metric measure of distance and therefore does not fully respect the ordinal nature of the data but is a pragmatic alternative to the widely used Euclidean distance which squares absolute differences. Feature vectors contributed by the same participant were treated as independent. The implementation of the *k*-medoids algorithm used to cluster the feature vectors can be derived as follows. A *k*-medoids algorithm randomly

assigns a user-defined '*k*' feature vectors to be the cluster centres (or medoids) and then iteratively (1) assigns each feature vector to the closest medoid and (2) recalculates the medoid of the clusters. The term 'medoid' refers to the use of actual data points as the centres for the clusters [35]. Such use of observed data as centres for the clusters prevents outputs such as "pain severity of 3.2", that might arise if means are used and that are uninterpretable and erroneously assume an interval scale. To implement the *k*-medoids algorithm, the Clustering Large Applications (CLARA) program was used, which was specifically designed for use of large data sets to reduce overall computation time [36, 37].

A *k*-medoids algorithm requires a user-defined value for the number of clusters (*k*) in the data [35]. The implementation of CLARA was therefore repeated for values of *k* from 1 to 20. The output of the algorithm can be sensitive to the random feature vectors selected as the medoids in the first stage of the algorithm. The algorithm was therefore repeated 20 times, once for each value of *k*. At each iteration, the remaining variance within each cluster was calculated as the 'within-cluster sum of squares' (WSS). The WSS calculates the total remaining distance between pairs of feature vectors in the same cluster. For each value of *k*, the iteration which returns the smallest value of WSS is selected and reported on a plot. From the plot of *k* against WSS, the optimal number of clusters can be chosen visually using the elbow method [38]. While ideally a formal trade-off would be made between model complexity and goodness of fit, there is no clear method to use. Existing methods (e.g. information criteria, silhouette method, gap statistic) can suggest different numbers of clusters [39], possibly due to under-penalising the complexity for datasets of the size used in this study. Therefore, the less formal elbow method allows us to be more explicit in the judgements we make, to resolve the absence of an unambiguous method for learning cluster numbers from data. To enable further analysis of clusters, and similarly to previous studies [e.g. 10, 30, 36–39], clusters were required to contain 5% of the trajectories.

Sensitivity analyses

Three sensitivity analyses were conducted to test assumptions made in the main analysis, with the methodology behind choosing these being to modify assumptions made about the data and to see if broad conclusions were robust. Robust conclusions would be indicative of a strong model-independent signal in the data, even if modified assumptions led to a less interpretable output. The main analysis assumed that data were on an ordered scale but relaxed the assumption that the data were longitudinal.

The first sensitivity analysis maintained the longitudinal nature of the data but implicitly assumed that the outcome variable was on a continuous scale. Feature vectors were compared using the Euclidean distance, which erroneously assumes regular intervals between values on the pain severity scale. However, the use of Euclidean distance permits the use of the *KmL* package, which specifically clusters longitudinal data [44]. The *KmL* package is an adaptation of the *k*-means algorithm. The *k*-means algorithm is similar to *k*-medoids, but the centre of each cluster is calculated using the mean of the feature vectors assigned to the cluster. The use of the *KmL* package, instead of the CLARA program, and the resulting use of mean trajectories rather than medoid trajectories, were the only adaptions to this sensitivity analysis. The feature vectors, the 20 repetitions of the algorithm for each value of *k*, and the use of elbow method to select *k* remained unchanged.

The second sensitivity analysis relaxed assumptions about both the longitudinal and ordinal nature of the outcome variable. In this sensitivity analysis, the data were not assumed to be longitudinal, and the outcome variable was assumed to be unordered categorical data. A different feature vector was used that converted ordinal pain-severity values into dummy variables using one-hot encoding. In this encoding, there were 35 binary categories, each representing a unique day and pain-severity category. The feature was recorded as ‘1’ if the pain-severity score was seen on that day, and ‘0’ otherwise. In this way, seven of the features were ‘1’ for each complete participant-week. The feature vectors were compared using the

Jaccard distance typically used for such vectors of binary data. The cluster analysis was then conducted using the CLARA program, in the same manner as described in the main analysis.

The third sensitivity analysis challenged the definition of a Monday–Sunday week when defining complete participant-weeks. Instead, the following analysis was conducted for each of the days (*D*) in the week. Complete participant-weeks were selected from the original data for each participant when there were pain severity data for each day in the *D* to *D*+6 week (e.g. Wednesday to Tuesday week). On each new dataset corresponding to a different *D*, clustering was conducted using the CLARA program at each value of *k* between 1 and 20, as described in the main analysis. Due to the adapted complete participant-weeks, individuals may have contributed different numbers of weeks to the sensitivity analysis.

For each sensitivity analysis, the optimal number of clusters was calculated. Similar numbers and descriptions of the clusters would provide evidence that the conclusions from the main analysis are robust. Furthermore, for each cluster in the main analysis, the proportion of trajectories allocated to the same cluster in each sensitivity analysis was calculated. A high proportion would further suggest that the results are robust to the assumptions made by using the Manhattan distance and *k*-medoids algorithm in the main analysis.

Description of clusters

Information about the trajectories assigned to each of the clusters in the optimal solution were summarised. First, the number of trajectories assigned to each cluster was reported. Secondly, the clusters were visualised with a spaghetti plot of individual trajectories and the medoid of each cluster. Finally, the average proportion of time spent in each cluster by each participant was calculated. This information was summarised by calculating the mean proportion of time spent in each cluster across demographics (i.e. age, sex, chronic-pain condition or conditions).

Transition between clusters

For the optimal solution of clusters in the main analysis, transition of individuals between clusters on consecutive weeks was examined. To do this, a subset of the total data was used. Complete participant-weeks ("This week") were retained if the participant had also contributed

a complete participant-week in the directly preceding week ("Last week"). A trajectory could be labelled as both "This week" and "Last week" if there were both preceding and succeeding weeks for the individual (Figure 3). The demographic data of participants included in this transition analysis were compared to those included in the main analysis.

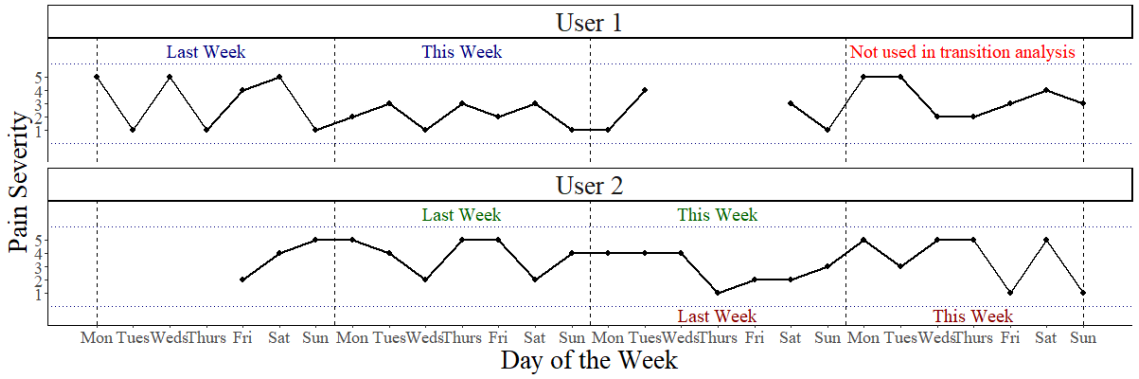


Figure 3: Example data from two participants, highlighting how their data is used to examine transitions between clusters. User 1 provided data in three complete participant-weeks. The first two are consecutive and therefore are used in the transition analysis. The final complete participant-week is not used. User 2 provided three complete participant-weeks. All three are used in the transition analysis. The middle week is labelled as both 'This Week' and 'Last Week' in different pairings.

Each trajectory was assigned a cluster in the CLARA program of the *k*-medoids cluster algorithm. The transition probabilities were then calculated as follows. For all trajectories in each cluster "Last week", the percentage of trajectories that transitioned to each cluster "This week" were calculated. The resulting percentages are reported in a transition matrix.

Data were analysed in R (version 4.1.2). Reporting of the analysis followed the STROBE guidelines [45].

Results

Data source

There were 2807 participants who contributed 21,919 participant-weeks of data to this analysis. The mean age was 51.2 (s.d. 12.8) years, and 83% were female. Table 1 reports the number of participants by age, sex, chronic pain condition and the average number of participant-weeks contributed to the analysis by members of the subgroup. Overall, older participants contributed a greater number of participant-weeks than younger participants. Men contributed slightly more (8.1) participant-weeks than women (7.7). Participants with osteoarthritis (9.1) and unspecific arthritis (9.0) contributed the highest number of participant-weeks, and participants with chronic headache (6.0) contributed the fewest participant-weeks. Comorbid conditions are described in Supplementary Materials 1, Table S1 and Table S2.

Table 1: Demographic information of the participants who contributed to the analysis and the average number of participant-weeks contributed by each subgroup.

		Number of participants (%)	Mean number of weekly trajectories contributed by participants
Age group (years)			
	17–24	67 (2.4)	5.2
	25–34	255 (9.1)	5.6
	35–44	508 (18.1)	6.8
	45–54	755 (26.9)	7.5
	55–64	788 (28.1)	8.6
	65–86	434 (15.5)	9.9
Sex			
	Female	2333 (83.1)	7.7
	Male	474 (16.9)	8.1
Chronic pain condition[a]			
	Rheumatoid arthritis	548 (19.5)	7.7
	Osteoarthritis	975 (34.7)	9.1
	Spondyloarthropathy	254 (9.0)	7.6
	Gout	96 (3.4)	7.8

	Unspecific arthritis	1028 (36.6)	9.0
	Fibromyalgia	718 (25.6)	7.1
	Chronic headache	274 (9.8)	6.0
	Neuropathic pain	427 (15.2)	7.5
	Other/no medical diagnosis	668 (23.8)	6.9

[a]Percentages exceed 100% because participants could report multiple chronic-pain conditions.

Identifying optimal number of clusters

The results of the CLARA algorithm are shown below. Figure 4 reports the remaining variability within clusters as the "within-cluster sum of squares" at each value of *k*. There is an elbow at $k = 4$, suggesting that most of the observed variability can be explained by a solution with four clusters with diminishing returns for including further clusters in the solution. Four clusters reduced the within-cluster sum of squares from 159,100 to 66,507 and so the clustering algorithm describes 58.2% percent of the variability in the data. Each cluster contained more than 5% of the pain trajectories. Therefore, four clusters provide an appropriate choice for these data.

The trajectories in each cluster are shown by the spaghetti plot in Figure 5. Trajectories are weighted such that thicker lines represent a higher number of trajectories following the path. The red line represents the medoid of the *k*-medoids algorithm. The clusters can be named by examining the medoid. They represent A = "no or low pain", B = "mild pain", C = "moderate pain", D = "severe pain". Cluster A contained 1714 (7.8%) trajectories, Cluster B contained 8246 (37.6%) trajectories, Cluster C contained 8376 (38.2%) trajectories, and Cluster D contained 3583 (16.3%) trajectories.

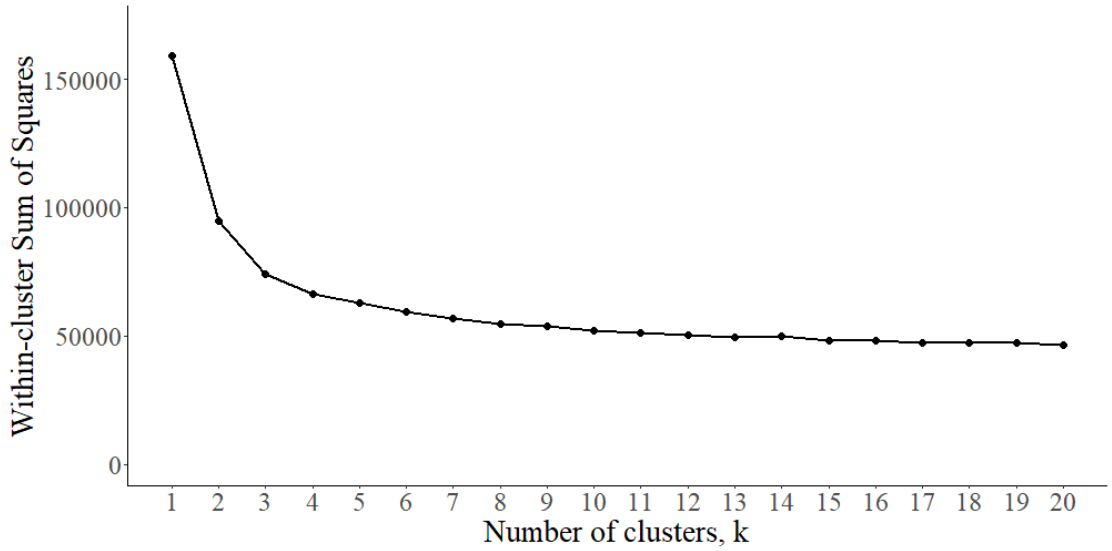


Figure 4: Unexplained variability across different cluster (*k*) solutions. The within-sum-of-squares indicates the remaining variance within clusters. An elbow at $k = 4$ suggests an appropriate solution, with diminishing returns for the inclusion of further clusters.

Sensitivity analyses

Three sensitivity analyses were conducted. First, trajectories were compared using Euclidean distances and clustered using the *KmL* algorithm. Full results are presented in Supplementary Materials 2. The plot visualising within-cluster sum of squares against *k* for this analysis is similar to that of the main analysis and has an elbow at $k = 4$ (Figure 6). The optimal four-cluster solution describes 60.0% of the observed variability. The descriptions of the spaghetti plots (i.e. Cluster A = no or low pain, Cluster B = mild pain etc) are the same as the main analysis, despite use of a mean rather than medoid to describe the average trajectory in each cluster. In total, 86.2% of trajectories were assigned to the same cluster as the main analysis, indicating similar results. Clusters B and C remain the largest clusters (38.7% and 36.5% of trajectories, respectively), although Cluster A is larger in this sensitivity analysis than the main analysis (11.4% in contrast to 7.8%).

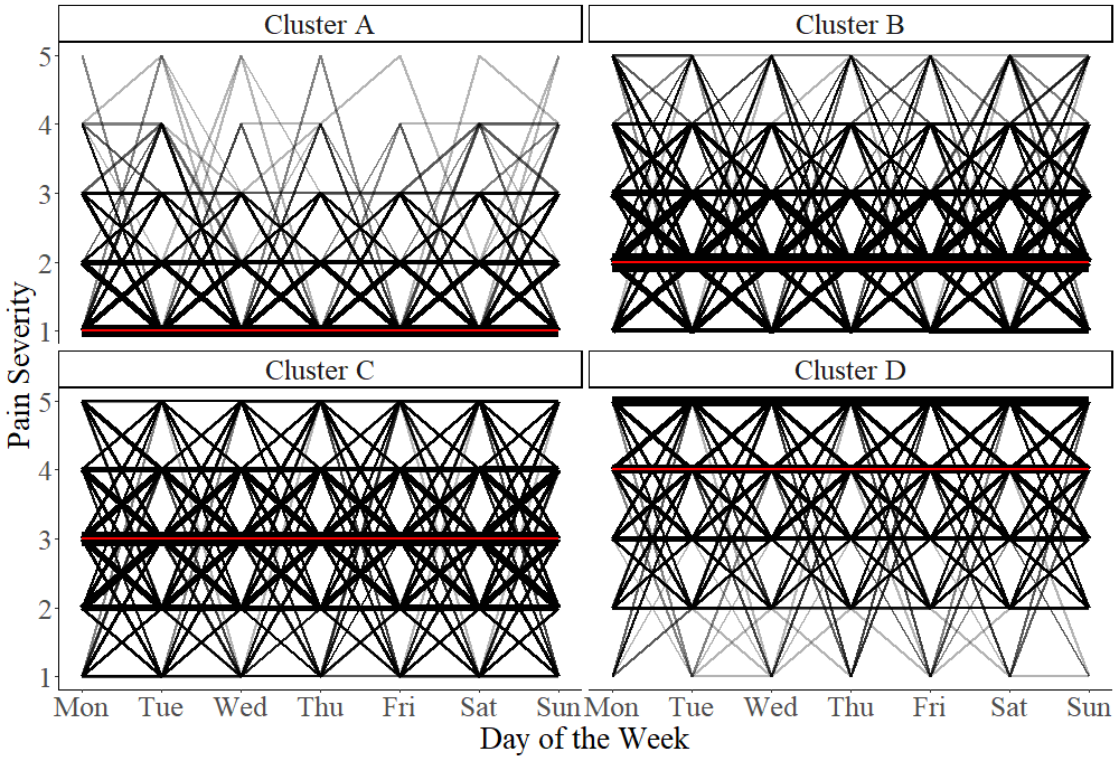


Figure 5: Weighted spaghetti plot of trajectories assigned to each cluster. The weight (and transparency) of each path represents the number of trajectories following that path. The red line represents the medoid of the cluster. Cluster A = "no or low pain", B = "mild pain", C = "moderate pain" and D = "severe pain".

The second sensitivity analysis compared one-hot trajectories using the Jaccard distance and clustered them using the CLARA program of the *k*-medoid algorithm. Full results are presented in Supplementary Materials 2. The plot of *k* against WSS for this analysis has an elbow at $k = 5$ (Figure 6). However, one cluster contained 4.5% of trajectories in the five-cluster solution, which did not meet the criteria for cluster sizes >5% and therefore a four-cluster solution remained optimal in this analysis. A four-cluster solution describes 50.0% of the variability.

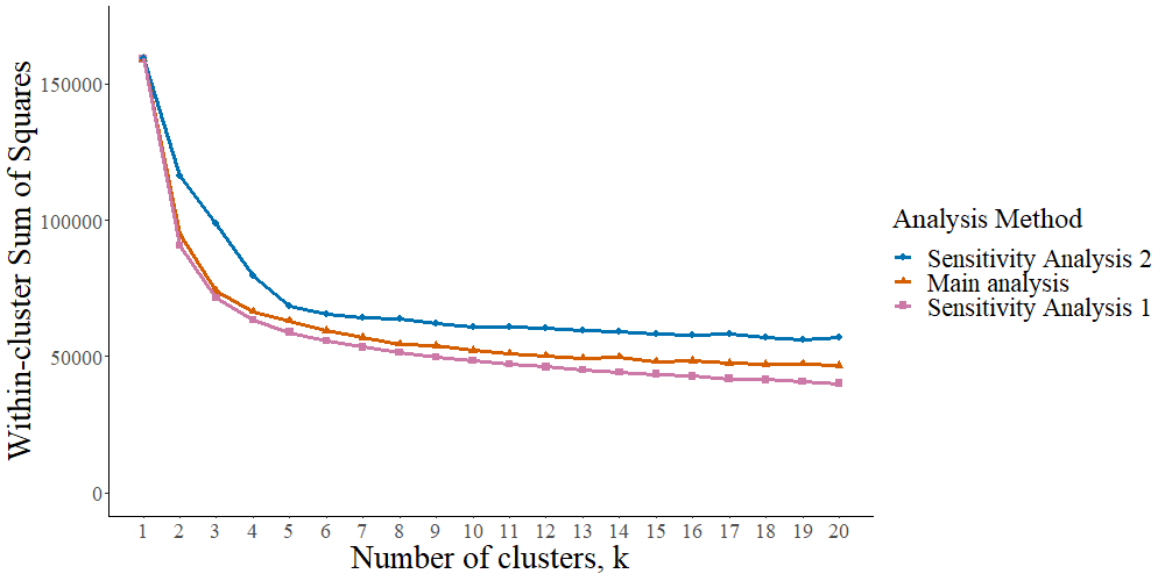


Figure 6: Unexplained variability across different cluster (*k*) solutions for the main analysis and two sensitivity analyses. In the main analysis and Sensitivity Analysis 1, there is an elbow at $k = 4$. In Sensitivity Analysis 2, there is an elbow at $k = 5$. Separate graphs are provided in Supplementary Materials 2.

Spaghetti plots of the four-cluster solution show the same descriptions as the main analysis. In total, 92.1% of trajectories were assigned to the same cluster as in the main analysis.

The third sensitivity analysis repeated the main analysis for different definitions of participant-weeks. Six sub-analyses were conducted, with each defining a participant-week beginning on a different day of the week. Full results are presented in Supplementary Materials 3. Each plot of within-cluster sum of squares against *k* suggested an optimal solution at $k = 4$. The proportions of trajectories assigned to each cluster in each four-cluster solution are similar to the main analysis. The proportions in Cluster A ranged between 7.7% and 7.9%, Cluster B between 37.5% and 37.7%, Cluster C between 38.0% and 38.4% and Cluster D between 16.2% and

16.4%. These results show that the main analysis is robust to the day of the week that the trajectories begin.

Description of clusters

The average proportion of time spent in different clusters across different characteristics (age, sex and condition) are summarised in Table 2. The oldest age bracket (65–86 years old) spent less time (9.8%) in the severe-pain cluster compared to the youngest age bracket (17–24, 28.3%). Women spent more time in the severe-pain cluster (18.0%) than men (12.3%) and less time in the lowest-pain cluster (Female: 6.5%, Male: 7.9%). Participants with fibromyalgia and neuropathic pain spent the most time in the severe-pain cluster (31.5% and 31.1%, respectively). Participants with rheumatoid arthritis spent the most time in the lowest-pain cluster (7.8%).

Transition between clusters

There were 12,267 pairs of participant weeks from 1761 participants used in the transition analysis. The demographic data are compared to those in the main analysis in Supplementary Materials 4. In general, a slightly higher proportion of older people contributed to the transition analysis than did to the cluster analysis. For example, people aged 65–86 contributed 15.5% of trajectories to the main analysis but 17.0% of trajectories to the transition analysis. There are no other differences in the demographics of participants contributing to the main analysis and the transition analysis.

The percentages of consecutive trajectories transitioning between clusters are shown in Figure 7. For each cluster, the highest percentage of trajectories in the consecutive week remain in the same cluster, with the percentage values ranging between 63% and 70%. On average, 66% of trajectories remain in the same cluster. When individuals move between clusters, it is most frequently to an adjacent cluster. There are a very small percentage of consecutive weeks displaying movement between clusters two or more levels away.

Table 2: Percentage of time spent in each cluster by baseline characteristics.

		Average percentage of time spent in Cluster A (%)	Average percentage of time spent in Cluster B (%)	Average percentage of time spent in Cluster C (%)	Average percentage of time spent in Cluster D (%)
All					
		6.7	36.8	39.4	17.0
Age group (years)					
	17–24	3.4	28.5	39.8	28.3
	25–34	6.7	32.0	41.5	19.8
	35–44	5.3	31.5	38.6	24.5
	45–54	5.6	36.3	39.4	18.7
	55–64	8.4	38.5	40.4	12.7
	65–86	7.9	44.9	37.3	9.8
Sex					
	Female	6.5	35.6	39.9	18.0
	Male	7.9	42.5	37.2	12.3
Chronic pain condition[a]					
	Rheumatoid arthritis	7.8	38.9	39.5	13.8
	Osteoarthritis	5.4	34.7	42.6	17.2
	Spondyloarthropathy	4.1	31.8	43.5	20.6
	Gout	6.1	33.0	41.6	19.3
	Unspecific arthritis	6.3	39.0	38.5	16.2
	Fibromyalgia	1.7	19.1	47.7	31.5
	Chronic headache	5.1	28.7	40.3	25.9
	Neuropathic pain	3.3	23.6	42.0	31.1
	Other/no medical diagnosis	6.7	36.8	39.4	17.0

For each characteristic, the average percentage of time spent in each cluster by members of the characteristic are reported. [a]Percentages exceed 100% because participants could report multiple chronic-pain conditions.

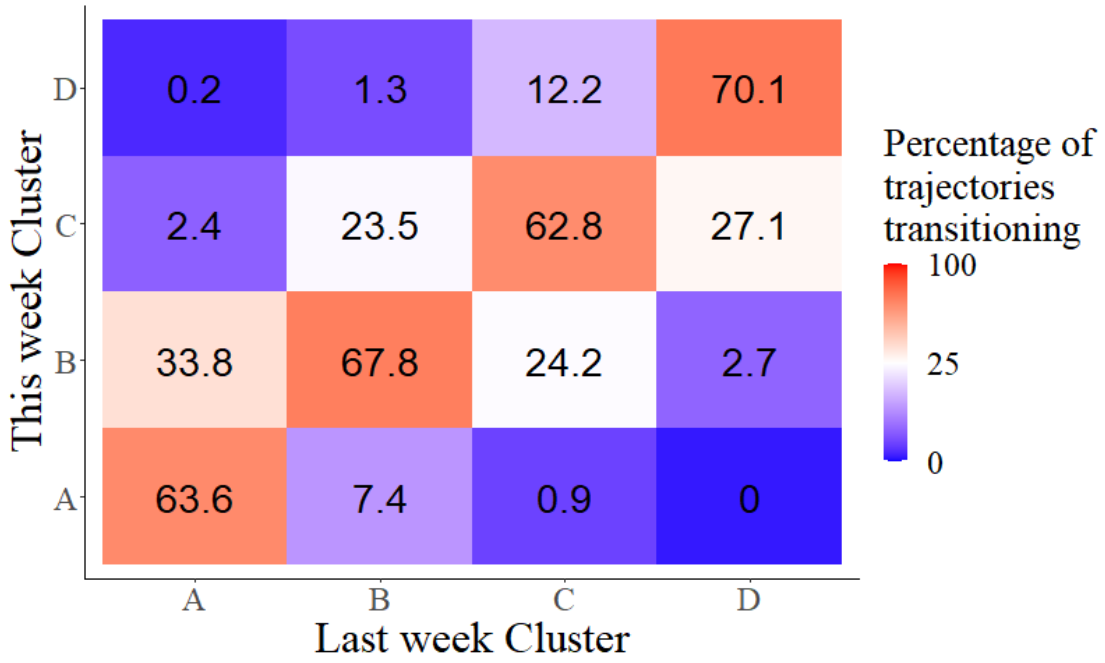


Figure 7: Transition matrix of movement between clusters on consecutive weeks. For membership in each cluster "Last Week", the percentage of membership in each cluster "This Week" is shown. For each combination, the expected percentage of transition would be 25%. Lower than expected percentage of membership (< 25%) is shown by blue boxes and higher than expected percentage of membership (> 25%) is shown by red boxes. The expected percentage (25%) is shown by white boxes, and values very close to 25% will appear white.

Discussion

Principal Findings

This study identified and described clusters of weekly trajectories of pain severity in a large population-based mHealth study, to address four objectives in turn. First, we reported that four clusters (A = "no/low pain severity", B = "mild pain", C = "moderate pain" and D = "severe pain") represented an optimal clustering solution for these data. In this solution, clusters B and C contained the greatest number of weekly pain trajectories.

Second, we conducted sensitivity analyses, to identify whether conclusions made about the first objective were robust to modified assumptions around the structure of the data. Two sensitivity analyses were conducted when the outcome variable was assumed to be (1) continuous and longitudinal and (2) categorical and not longitudinal. These analyses found that four clusters remained a suitable conclusion. A third sensitivity analysis found no differences in the clusters of trajectories starting on different days of the week.

Third, younger people and women contributed a greater number of trajectories to the severe-pain cluster when compared to older people and men, respectively. Participants with fibromyalgia and neuropathic pain contributed more trajectories to the severe-pain cluster than participants with other pain conditions, whereas participants with rheumatoid arthritis contributed more trajectories of to the no/low pain cluster than participants with other pain conditions.

Fourth, we examined transitions between clusters and found that about 66% of consecutive trajectories contributed to the same cluster. However, there was clear evidence of between-cluster movement with 34% of consecutive trajectories assigned to different clusters. Between-cluster movement was most likely to a neighbouring cluster; for example, moving from cluster 1 to cluster 2 was more common than moving from cluster 1 to cluster 3. This analysis demonstrates that overall, individuals tend to experience similar patterns of pain severity from week to week, although there are substantial experiences of increase or decreases in pain severity, thereby reflecting the lived experience of people with chronic pain having variability in symptoms, and noting how pain can fluctuate between weeks..

People with chronic pain have highlighted a need to describe and predict the variability in the severity of their pain. Through clustering, this study has described four common experiences of pain severity, accounting for two-thirds of the observed variability. However, trajectories within each cluster are not homogenous and there remains within-cluster variation. To better describe the individual weekly pain experience, future work should explore the remaining variability within clusters.

Comparison with Prior Work

Many studies have identified clusters of pain trajectories among individuals living with chronic pain. Some have focussed on participants with one chronic pain condition, such as osteoarthritis [13, 30–33, 46–52], low back pain [11, 12, 22, 53–59], other back pain [20, 27, 43, 60], neck or shoulder pain [28, 55, 61, 62], leg pain [24], knee pain [63] or foot pain [64], whereas others have identified clusters among a broader population such as those with musculoskeletal pain [21, 26, 41, 65, 66] or general pain [42, 67–69]. Clusters in these studies were described by the severity of pain (e.g. no pain, very low pain, mild pain, moderate pain, high pain, severe pain), the level of change in pain severity (e.g. persistent, ongoing, episodic, worsening, recovering, fluctuating) or a combination of these features.

These previous studies have often considered only sparse data, with relatively large time intervals between consecutive data points. Of those gathering data for at least one year (27 studies), data were collected more than twelve times in only two studies [11, 61]. In these two studies, data were collected weekly to explore the course of specific pain conditions (neck pain, low back pain). The clusters identified in our analysis were described by the severity of pain, similar to clusters in studies with sparse data. Our clusters were unlikely to identify long-term disease development, as with trajectories over longer periods.

In the listed studies, the experiences of individuals were described by a single trajectory across the full duration of follow-up, whereas our study examined week-to-week transition between clusters. Kongsted et al. [70] have previously examined week-to-week pain severity across a year, using pre-defined clusters. They identified that 41% and 21% of respondents in two different datasets had stable pain over a year, defined as pain within 1 point of the mean pain value on an 11-point numerical rating scale. The remaining pain trajectories were classified as having a single episode of pain, being episodic or fluctuating. The transitions identified in our study suggest stability between 66% of consecutive weeks. However, some individuals in our study may experience the other longer-term descriptions from Kongsted et al. For example, an individual might not transition out of a cluster for most of the year yet experience a single

episode. Future studies should further examine the movement between different pain states and identify the drivers of these transitions.

Strengths and Limitations

A number of strengths and limitations of this study should be considered. First, a strength was that participants could contribute daily data for up to 64 weeks. This frequent and granular data collection, enabled by mHealth, overcame limitations of sparse data collection in previous studies (as identified by Beukenhorst et al. [71]). As a result, this study was able to analyse the weekly trajectories contributed by participants, determining common pain patterns among a chronic pain population at a more granular scale than previously investigated.

Second, the analysis presented in this paper modelled weekly pain trajectories rather than individual people. In contrast to studies that assign each individual to a single cluster across the whole follow-up, individuals were able to transition between different pain clusters over time as their pain experience changed and their condition developed. These transitions were observed in 34% of consecutive weeks, and this flexibility can be used in future work to explore the mechanisms driving movement between clusters.

Third, assumptions about the ordinal, longitudinal form of the data were modified in sensitivity analyses. A four-cluster solution was most suitable for each analysis, indicating a strong model-independent signal in the data and a more robust conclusion about the most suitable number of clusters. Furthermore, the assignment of trajectories to each cluster were similar in each analysis (at least 86% similarity), indicating further stability in the results. There were benefits to the use of both the *k*-medoids algorithm and longitudinal adaptation of the *k*-means algorithm used in this analysis. First, neither of these methods require parametric assumptions about the form of the data [44]. Second, no prior assumptions, including the shape of the trajectory, are required by the algorithms [72]. Therefore, this data-driven approach made limited assumptions about the form of the data.

There were also limitations to the study. The data used in this study were from a population-based study which represented the UK population. Cloudy with a Chance of Pain recruited participants from all UK postcodes, although men and those in the age brackets 17–34 and 75+ were underrepresented in the study population [19]. Despite being a smaller population, older people and men contributed more trajectories on average and were more likely to contribute trajectories to a less severe pain category. As these clusters will be used in the development of a pain-forecasting model, clusters should be generalisable to the chronic-pain population, and there remains the possibility that different pain clusters and between-cluster transitions could be realized among those who did and did not contribute to the study. Although it is unlikely that our large study population would display different pain clusters and transitions to the chronic-pain population that would use a smartphone tracking application, it remains a possibility that should be explored in future studies.

This analysis further selected participants by the requirement to provide a week of complete pain-severity data, thus excluding missing data. There are reasons that data might be missing not at random, including missing due to severe pain, missing due to low pain severity, and missing due to stable pain that results in repetitive score input and thus disengagement. The transition analysis also further selected participants by requiring two weeks of complete pain-severity data. However, the age, sex and chronic-pain conditions of respondents in the main analysis and transition analysis (Table 4) were similar to those in the full-study population (Supplementary Table 1 in Dixon et al. [14]), suggesting that the included participants were representative of the study population.

There were limitations in the method used for clustering. First, the absence of parametric assumptions in either the *k*-medoids algorithm or KmL package resulted in goodness-of-fit measures being inappropriate [72]. Therefore, the elbow method was used to select the optimum number of clusters. However, the use of the elbow method introduces subjectivity. Second, both the *k*-medoids algorithm and the KmL package require random starting values for the cluster centres, which can add volatility to the results. This volatility was mitigated by repeating the

algorithms 20 times each and selecting the solution with the lowest remaining variability within clusters.

Conclusions

Previous research has highlighted a need to better understand pain variability experienced by individuals with chronic pain [9]. Feelings of uncertainty among people with chronic pain have led them to want to better understand the pain that they may experience in the future. Clustering weekly pain trajectories offers a first step to better understanding common experiences of pain severity. Once these common experiences are better described, they can be used in future work to predict movement between clusters.

There are limited methods available for clustering pain severity that respect the ordinal and longitudinal nature of patient-generated health data in a computationally inexpensive manner. The clustering method and subsequent sensitivity analyses presented in this paper suggest that the use of *k*-medoids is robust to assumptions about the data structure.

This study has identified four distinct patterns of weekly pain severity: no or low pain, mild pain, moderate pain or severe pain. These can be used to describe short-term pain experiences of people with chronic pain. Future work is required to identify how these clusters can be used in a forecasting model of pain. First, there remains individual variability within clusters of pain severity. Participants of PPI have identified that fluctuations in pain severity should be forecasted, and therefore within-cluster variability should be quantified to further understand the weekly pain experience of individuals. Second, transition of individuals between clusters should be explored to identify drivers of movement between pain clusters on an individual level. The clusters identified in this study and future work to understand within-cluster variability and drivers of movement between clusters will enable a future pain-forecasting model.

Conflicts of interest:

WGD has received consultancy fees from Google, and DMS has received consultancy fees from Palta, both unrelated to this work. All other authors have no conflicts of interest to declare.

Funding:

This study was supported by the Centre for Epidemiology Versus Arthritis (grant number 21755). Cloudy with a Chance of Pain was funded by Versus Arthritis (grant reference 21225), with additional support from the Centre for Epidemiology (grants 21755 and 20380). TH receives funding from the Royal Society (grant number INF/R2/180067) and the Alan Turing Institute for Data Science and Artificial Intelligence. The funders of the study had no role in study design, data interpretation, writing of the manuscript, or the decision to submit for publication.

Data availability statement:

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Code Availability:

Data management and analyses were performed in R 4.1.2. Code is available on reasonable request.

Reference List

1. McQueenie R, Jani BD, Siebert S, McLoone P, McCowan C, Macdonald S, et al. Prevalence of chronic pain in LTCs and multimorbidity: A cross-sectional study using UK Biobank. J Multimorb Comorbidity 2021;11. doi:10.1177/26335565211005870

2. Wolfe F, Clauw DJ, Fitzcharles M-A, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res 2010;62(5):600–610. doi:10.1002/acr.20140

3. Steineck A, Bradford MC, O'Daffer A, Fladeboe KM, O'Donnell MB, Scott S, et al. Quality of Life in Adolescents and Young Adults: The Role of Symptom Burden. J Pain Symptom Manage 2022 Sep;64(3):244–253.e2. doi:10.1016/j.jpainsymman.2022.05.017

4. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005 Jan;113(1–2):9–19. doi:10.1016/j.pain.2004.09.012

5. Crowe M, Whitehead L, Seaton P, Jordan J, Mccall C, Maskill V, Trip H. Qualitative meta-synthesis: the experience of chronic pain across conditions. J Adv Nurs 2017 May;73(5):1004–1016. doi:10.1111/jan.13174

6. Schneider S, Junghaenel DU, Keefe FJ, Schwartz JE, Stone AA, Broderick JE. Individual differences in the day-to-day variability of pain, fatigue, and well-being in patients with rheumatic disease: associations with psychological variables. Pain 2012 Apr;153(4):813–822. doi:10.1016/j.pain.2012.01.001

7. Wesolowicz DM, Bishop MD, Robinson ME. An Examination of Day-to-Day and Intraindividual Pain Variability in Low Back Pain. Pain Med 2021 Oct;22(10):2263–2275. doi:10.1093/pm/pnab119

8. Johnson LM, Zautra AJ, Davis MC. The role of illness uncertainty on coping with fibromyalgia symptoms. Health Psychol 2006 Nov;25(6):696–703. doi:10.1037/0278-6133.25.6.696

9. Little CL, Druce KL, Dixon WG, Schultz DM, House T, McBeth J. What do people living with chronic pain want from a pain forecast? A research prioritisation study. medRxiv 2023 Apr; doi:10.1101/2023.04.24.23289032

10. Flurey CA, Morris M, Richards P, Hughes R, Hewlett S. It's like a juggling act: rheumatoid arthritis patient perspectives on daily life and flare while on current treatment regimes. Rheumatology 2014;53(4):696–703. doi:10.1093/rheumatology/ket416

11. Kongsted A, Kent P, Hestbaek L, Vach W. Patients with low back pain had distinct clinical course patterns that were typically neither complete recovery nor constant pain. A latent class analysis of longitudinal data. Spine J 2015 May;15(5):885–894. doi:10.1016/j.spinee.2015.02.012

12. Dunn KM, Jordan K, Croft PR. Characterizing the Course of Low Back Pain: A Latent Class Analysis. Am J Epidemiol 2006 Apr;163(8):754–761. doi:10.1093/aje/kwj100

13. Bastick AN, Wesseling J, Damen J, Verkleij SPJ, Emans PJ, Bindels PJE, Bierma-Zeinstra SMA. Defining knee pain trajectories in early symptomatic knee osteoarthritis in primary care: 5-year results from a nationwide prospective cohort study (CHECK). Br J Gen Pract 2016 Jan;66(642):e32-39. doi:10.3399/bjgp15X688129

14. Dixon WG, Beukenhorst AL, Yimer BB, Cook L, Gasparrini A, El-Hay T, et al. How the weather affects the pain of citizen scientists using a smartphone app. NPJ Digit Med 2019 Oct;2:105. doi:10.1038/s41746-019-0180-3

15. Check Hayden E. Mobile-phone health apps deliver data bounty. Nature 2016 Mar;531(7595):422–423. doi:10.1038/531422a

16. Shiffman S, Stone AA, Hufford MR. Ecological Momentary Assessment. Annu Rev Clin Psychol 2008;4:1–32. doi:10.1146/annurev.clinpsy.3.022806.091415

17. Liddell TM, Kruschke JK. Analyzing ordinal data with metric models: What could possibly go wrong? J Exp Soc Psychol 2018;79:328–348. doi:10.1016/j.jesp.2018.08.009

18. Schultz DM, Beukenhorst AL, Yimer BB, Cook L, Pisaniello HL, House T, et al. Weather Patterns Associated with Pain in Chronic-Pain Sufferers. Bull Am Meteorol Soc 2020;101(5):E555–E566. doi:10.1175/BAMS-D-19-0265.1

19. Druce KL, McBeth J, van der Veer SN, Selby DA, Vidgen B, Georgatzis K, et al. Recruitment and Ongoing Engagement in a UK Smartphone Study Examining the Association Between Weather and Pain: Cohort Study. JMIR MHealth UHealth 2017 Nov;5(11):e168. doi:10.2196/mhealth.8162

20. Knecht C, Hartnack S, Sick B, Riner F, Schweinhardt P, Wirth B. A prospective observational study on trajectories and prognostic factors of mid back pain. BMC Musculoskelet Disord 2020 Aug;21(1):554. doi:10.1186/s12891-020-03534-5

21. Crook J, Moldofsky H. The clinical course of musculoskeletal pain in empirically derived groupings of injured workers. Pain 1996 Oct;67(2–3):427–433. doi:10.1016/0304-3959(96)03151-x

22. Axén I, Bodin L, Bergström G, Halasz L, Lange F, Lövgren PW, et al. Clustering patients on the basis of their individual course of low back pain over a six month period. BMC Musculoskelet Disord 2011 May;12:99. doi: 10.1186/1471-2474-12-99

23. Brennan PL, SooHoo S. Effects of Mental Health Disorders on Nursing Home Residents' Nine-Month Pain Trajectories. Pain Med 2020 Mar;21(3):488–500. doi:10.1093/pm/pnz177

24. Ogollah RO, Konstantinou K, Stynes S, Dunn KM. Determining One-Year Trajectories of Low-Back-Related Leg Pain in Primary Care Patients: Growth Mixture Modeling of a Prospective Cohort Study. Arthritis Care Res 2018 Dec;70(12):1840–1848. doi:10.1002/acr.23556

25. Battaglia M, Garon-Carrier G, Brendgen M, Feng B, Dionne G, Vitaro F, et al. Trajectories of pain and anxiety in a longitudinal cohort of adolescent twins. Depress Anxiety 2020 May;37(5):475–484. doi:10.1002/da.22992

26. Neupane S, Nygård C-H, Prakash KC, von Bonsdorff MB, von Bonsdorff ME, Seitsamo J, et al. Multisite musculoskeletal pain trajectories from midlife to old age: a 28-year follow-up of municipal employees. Occup Environ Med 2018 Dec;75(12):863–870. doi:10.1136/oemed-2018-105235

27. Dunn KM, Campbell P, Jordan KP. Long-term trajectories of back pain: cohort study with 7-year follow-up. BMJ Open 2013 Dec;3(12):e003838. doi:10.1136/bmjopen-2013-003838

28. Hallman DM, Rasmussen CDN, Jørgensen MB, Holtermann A. Time course of neck-shoulder pain among workers: A longitudinal latent class growth analysis. Scand J Work Environ Health 2018 Jan;44(1):47–57. doi:10.5271/sjweh.3690

29. Hebert JJ, Abraham E, Wedderkopp N, Bigney E, Richardson E, Darling M, et al. Patients undergoing surgery for lumbar spinal stenosis experience unique courses of pain and disability: A group-based trajectory analysis. PloS One 2019 Nov;14(11):e0224200. doi:10.1371/journal.pone.0224200

30. Wieczorek M, Rotonda C, Coste J, Pouchot J, Saraux A, Guillemin F, Rat A-C.Trajectory analysis combining pain and physical function in individuals with knee and hip osteoarthritis: results from the French KHOALA cohort. Rheumatol 2020 Nov;59(11):3488–3498. doi:10.1093/rheumatology/keaa148

31. Dai Z, Lu N, Niu J, Felson DT, Zhang Y. Dietary Fiber Intake in Relation to Knee Pain Trajectory. Arthritis Care Res 2017 Sep;69(9):1331–1339. doi:10.1002/acr.23158

32. Radojčić MR, Arden NK, Yang X, Strauss VY, Birrell F, Cooper C, et al. Pain trajectory defines knee osteoarthritis subgroups: a prospective observational study. Pain 2020 Dec;161(12):2841–2851. doi:10.1097/j.pain.0000000000001975

33. Pan F, Tian J, Aitken D, Cicuttini F, Jones G. Predictors of pain severity trajectory in older adults: a 10.7-year follow-up study. Osteoarthritis Cartilage 2018 Dec;26(12):1619–1626. doi:10.1016/j.joca.2018.08.002

34. Pan F, Tian J, Cicuttini F, Jones G. Metabolic syndrome and trajectory of knee pain in older adults. Osteoarthritis Cartilage 2020 Jan;28(1):45–52. doi:10.1016/j.joca.2019.05.030

35. Kassambara A. Practical guide to cluster analysis in R: unsupervised machine learning, Edition 1. United States: STHDA; 2017. ISBN:9781542462709

36. Kaufman L, Rousseeuw PJ. Clustering Large Data Sets. Pattern Recognition in Practice II: proceedings of an international workshop held in Amsterdam, June 19–21, 1985. Elsevier; 1986. p. 425–437. ISBN:9780444878779

37. Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. Vol.344. Hoboken, N.J: Wiley; 2005. ISBN:9780471735786

38. Sarang P. Centroid-Based Clustering: Clustering Algorithms for Hard Clustering. In: Thinking Data Science. Cham: Springer International Publishing; 2023. p. 171–183. ISBN:9783031023637

39. Hardy A. On the number of clusters. Comput Stat Data Anal 1996 Nov;23(1):83–96. doi:10.1016/S0167-9473(96)00022-9

40. Thomas MJ, Yu D, Nicholls E, Bierma-Zeinstra S, Conaghan PG, Stoner KJ, et al. Short-Term Recovery Trajectories of Acute Flares in Knee Pain: A UK-Netherlands Multicenter Prospective Cohort Analysis. Arthritis Care Res 2020 Dec;72(12):1687–1692. doi:10.1002/acr.24088

41. Airila A, Hakanen JJ, Luukkonen R, Lusa S, Punakallio A, Leino-Arjas P. Developmental trajectories of multisite musculoskeletal pain and depressive symptoms: the effects of job demands and resources and individual factors. Psychol Health 2014;29(12):1421–1441. doi:10.1080/08870446.2014.945929

42. James RJE, Walsh DA, Ferguson E. General and disease-specific pain trajectories as predictors of social and political outcomes in arthritis and cancer. BMC Med 2018 Apr;16(1):51. doi:10.1186/s12916-018-1031-9

43. Canizares M, Rampersaud YR, Badley EM. Course of Back Pain in the Canadian Population: Trajectories, Predictors, and Outcomes. Arthritis Care Res 2019 Dec;71(12):1660–1670. doi:10.1002/acr.23811

44. Genolini C, Falissard B. Kml: A package to cluster longitudinal data. Comput Methods Programs Biomed 2011 Dec;104(3):e112–e121. doi:10.1016/j.cmpb.2011.05.008

45. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Bull World Health Organ 2007 Nov;85(11):867–872. doi:10.2471/blt.07.045120

46. Verkleij SPJ, Hoekstra T, Rozendaal RM, Waarsing JH, Koes BW, Luijsterburg PAJ, Bierma-Zeinstra SMA. Defining discriminative pain trajectories in hip osteoarthritis over a 2-year time period. Ann Rheum Dis 2012 Sep;71(9):1517–1523. doi:10.1136/annrheumdis-2011-200687

47. Wesseling J, Bastick AN, ten Wolde S, Kloppenburg M, Lafeber FPJG, Bierma-Zeinstra SMA, Bijlsma JWJ. Identifying Trajectories of Pain Severity in Early Symptomatic Knee Osteoarthritis: A 5-year Followup of the Cohort Hip and Cohort Knee (CHECK) Study. J Rheumatol 2015 Aug;42(8):1470–1477. doi:10.3899/jrheum.141036

48. Nicholls E, Thomas E, van der Windt DA, Croft PR, Peat G. Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: findings from the Knee Clinical Assessment Study and the Osteoarthritis Initiative. Osteoarthritis Cartilage 2014 Dec;22(12):2041–2050. doi:10.1016/j.joca.2014.09.026

49. Bastick AN, Verkleij SPJ, Damen J, Wesseling J, Hilberdink WKHA, Bindels PJE, Bierma-Zeinstra SMA. Defining hip pain trajectories in early symptomatic hip osteoarthritis – 5 year results from a nationwide prospective cohort study (CHECK). Osteoarthritis Cartilage 2016 May;24(5):768–775. doi:10.1016/j.joca.2015.11.023

50. Collins JE, Katz JN, Dervan EE, Losina E. Trajectories and risk profiles of pain in persons with radiographic, symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage 2014 May;22(5):622–630. doi:10.1016/j.joca.2014.03.009

51. Johnson AJ, Vasilopoulos T, Booker SQ, Cardoso J, Terry EL, Powell-Roach K, et al. Knee pain trajectories over 18 months in non-Hispanic Black and non-Hispanic White adults with or at risk for knee osteoarthritis. BMC Musculoskelet Disord 2021 May;22(1):415. doi:10.1186/s12891-021-04284-8

52. Atukorala I, Downie A, Pathmeswaran A, Deveza LMA, Chang T, Zhang Y, Hunter DJ. Short-term pain trajectories in patients with knee osteoarthritis. Int J Rheum Dis 2022 Mar;25(3):281–294. doi:10.1111/1756-185X.14267

53. Tamcan O, Mannion AF, Eisenring C, Horisberger B, Elfering A, Müller U. The course of chronic and recurrent low back pain in the general population. Pain 2010 Sep;150(3):451–457. doi:10.1016/j.pain.2010.05.019

54. Coenen P, Smith A, Paananen M, O'Sullivan P, Beales D, Straker L. Trajectories of Low Back Pain From Adolescence to Young Adulthood. Arthritis Care Res 2017 Mar;69(3):403–412. doi:10.1002/acr.22949

55. Ailliet L, Rubinstein SM, Hoekstra T, van Tulder MW, de Vet HCW. Long-term trajectories of patients with neck pain and low back pain presenting to chiropractic care: A latent class growth analysis. Eur J Pain 2018 Jan;22(1):103–113. doi: 10.1002/ejp.1094

56. Downie AS, Hancock MJ, Rzewuska M, Williams CM, Lin C-WC, Maher CG. Trajectories of acute low back pain: a latent class growth analysis. Pain 2016 Jan;157(1):225–234. doi:10.1097/j.pain.0000000000000351

57. Panken G, Hoekstra T, Verhagen A, van Tulder M, Twisk J, Heymans MW. Predicting chronic low-back pain based on pain trajectories in patients in an occupational setting: an exploratory analysis. Scand J Work Environ Health 2016 Jun;42(6):520–527. doi:10.5271/sjweh.3584

58. Macedo LG, Maher CG, Latimer J, McAuley JH, Hodges PW, Rogers WT. Nature and determinants of the course of chronic low back pain over a 12-month period: a cluster analysis. Phys Ther 2014 Feb;94(2):210–221. doi:10.2522/ptj.20120416

59. Kyrönlahti SM, Nygård C-H, Prakash K C, Neupane S. Trajectories of low back pain from midlife to retirement and functional ability at old age. Eur J Public Health 2022 Jun;32(3):497–503. doi:10.1093/eurpub/ckab191

60. Enthoven WTM, Koes BW, Bierma-Zeinstra SMA, Bueving HJ, Bohnen AM, Peul WC, et al. Defining trajectories in older adults with back pain presenting in general practice. Age Ageing 2016 Nov;45(6):878–883. doi:10.1093/ageing/afw127

61. Pico-Espinosa OJ, Côté P, Hogg-Johnson S, Jensen I, Axén I, Holm LW, Skillgate E. Trajectories of Pain Intensity Over 1 Year in Adults With Disabling Subacute or Chronic Neck Pain. Clin J Pain 2019 Aug;35(8):678–685. doi:10.1097/AJP.0000000000000727

62. Eriks-Hoogland IE, Hoekstra T, de Groot S, Stucki G, Post MW, van der Woude LH. Trajectories of musculoskeletal shoulder pain after spinal cord injury: Identification and predictors. J Spinal Cord Med 2014 May;37(3):288–298. doi:10.1179/2045772313Y.0000000168

63. Haybatollahi SM, James RJE, Fernandes G, Valdes A, Doherty M, Zhang W, et al. Identifying multiple knee pain trajectories and the prediction of opioid and NSAID

medication used: A latent class growth approach. Pain Pract 2022 Feb;22(2):210–221. doi:10.1111/papr.13082

64. Marshall M, Blagojevic-Bucknall M, Rathod-Mistry T, Thomas MJ, Edwards JJ, Peat G, et al. Identifying Long-Term Trajectories of Foot Pain Severity and Potential Prognostic Factors: A Population-Based Cohort Study. Arthritis Care Res 2023 May;75(5). doi:10.1002/acr.24823

65. Leino-Arjas P, Rajaleid K, Mekuria G, Nummi T, Virtanen P, Hammarström A. Trajectories of musculoskeletal pain from adolescence to middle age: the role of early depressive symptoms, a 27-year follow-up of the Northern Swedish Cohort. Pain 2018 Jan;159(1):67–74. doi:10.1097/j.pain.0000000000001065

66. Oakman J, Neupane S, Kyrönlahti S, Nygård C-H, Lambert K. Musculoskeletal pain trajectories of employees working from home during the COVID-19 pandemic. Int Arch Occup Environ Health 2022 Nov;95(9):1891–1901. doi:10.1007/s00420-022-01885-1

67. Neupane S, Lallukka T, Pietiläinen O, Rahkonen O, Leino-Arjas P. Trajectories of multisite musculoskeletal pain in midlife: Associations with common mental disorders. Eur J Pain 2020 Feb;24(2):364–373. doi:10.1002/ejp.1492

68. Sun F, Park NS, Wardian J, Lee BS, Roff LL, Klemmack DL, et al. Predicting the Trajectories of Perceived Pain Intensity in Southern Community-Dwelling Older Adults: The Role of Religiousness. Res Aging 2013 Nov;35(6). doi:10.1177/0164027512456402

69. Wickrama KAS, Lee TK, O'Neal CW. Heterogeneity in Husbands' and Wives' Physical Pain Trajectories Over Mid-Later Years: Biopsychosocial Stratification and Implications for Later-Life Well-Being. The Gerontologist 2021 Aug;61(6):930–941. doi:10.1093/geront/gnab007

70. Kongsted A, Hestbæk L, Kent P. How can latent trajectories of back pain be translated into defined subgroups? BMC Musculoskelet Disord 2017 Jul;18(1):285. doi:10.1186/s12891-017-1644-8

71. Beukenhorst AL, Schultz DM, McBeth J, Sergeant JC, Dixon WG. Are weather conditions associated with chronic musculoskeletal pain? Review of results and methodologies. Pain 2020 Apr;161(4):668–683. doi:10.1097/j.pain.0000000000001776

72. Genolini C, Alacoque X, Sentenac M, Arnaud C. kml and kml3d: *R* Packages to Cluster Longitudinal Data. J Stat Softw 2015;65(4):1–34. doi:10.18637/jss.v065.i04

Abbreviations

App: application

mHealth: mobile health

s.d.: standard deviation

WSS: within-cluster sum of squares

5. Manuscript 3: Patterns and predictors of pain variability in patient-generated daily-symptom data: Results from a smartphone app study

Chapter 3 reported that people with chronic pain showed interest in understanding their pain fluctuations. Chapter 4 reported that 66% of pain variability could be attributed to between-person changes shown as the clusters. However, this means that 34% of pain variability is unaccounted for. The present chapter further explores this variability. It does so by describing the variability observed, using measures discussed in Chapter 2, and also by identifying covariates that are associated with changes in pain severity. As group-level associations may be obscured by population-level analysis, the work presented in this chapter was conducted within clusters.

I led meetings with all coauthors to discuss the conceptualization of this study. I worked with coauthor BBY to develop the methodology used in this work. I developed the code and conducted the statistical analysis described in this manuscript. I wrote the original draft of the manuscript. All coauthors contributed to reviewing and editing of the manuscript.

The following manuscript has been submitted for peer review at *Lancet Digital Health*. Supplementary materials are provided in Appendix F.

Patterns and predictors of pain variability in patient-generated daily-symptom data:

Results from a smartphone app study

Submission Type: Article

Authors:

Claire L Little, MA (Corresponding author)
Centre for Epidemiology Versus Arthritis,
University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
claire.little@postgrad.manchester.ac.uk

Dr Belay B Yimer, PhD
Centre for Epidemiology Versus Arthritis,
University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

Prof Thomas House, PhD
Department of Mathematics
University of Manchester, Upper Brook Street, Manchester, M13 9SR, United Kingdom

Prof William G Dixon, MRCP, PhD
Centre for Epidemiology Versus Arthritis,
University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
NIHR Manchester Musculoskeletal Biomedical Research Unit,
Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Prof David M Schultz, PhD
Department of Earth and Environmental Sciences,
Centre for Atmospheric Science,
University of Manchester, Oxford Road, Manchester, M13 9PS, United Kingdom
Centre for Crisis Studies and Mitigation
University of Manchester, Manchester, United Kingdom

Prof John McBeth, PhD
Centre for Epidemiology Versus Arthritis,
University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
NIHR Manchester Musculoskeletal Biomedical Research Unit,
Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Abstract

Background

Digital health technologies support the collection of patient-generated health data that is frequent, longitudinal, and collected in participant's own environments. Such high-frequency data could detect patterns of variation in disease and associated symptoms, although the methods for quantifying variability are unclear. Here, we explore variability in pain severity in people living with chronic-pain conditions. The study aims were to (1) quantify the amount of day-to-day variability in pain severity across a seven-day period and (2) identify variables associated with day-to-day variability in pain severity.

Methods

Data were collected via a smartphone application from a population-based mobile-health (mHealth) study, Cloudy with a Chance of Pain. Pain changes were calculated on consecutive days. Summary statistics and distributions of pain changes were calculated within complete weeks of data, which had previously been assigned to one of four clusters. Cumulative-probit models were used to identify associations between changes in pain severity and changes in exposure data collected in the mHealth study. These exposure data were collected actively (e.g., fatigue) and passively (e.g., weather).

Findings

A cluster of low/no-pain severity had a higher proportion of weeks with no within-week changes in pain severity than other clusters. Across all clusters, observed changes were most frequently of one unit (out of five), but larger fluctuations were also observed. Changes in pain severity were associated most strongly with changes in pain interference and were also associated with changes in other variables including fatigue, morning stiffness, mood, and participant well-being.

Interpretation

Data collected frequently through digital-health tools can be used to explore variability in symptoms. Pain interference was most strongly associated to changes in pain severity but changes in modifiable variables (fatigue and mood) were associated with changes in pain severity; therefore, these may be the foci of future work.

Funding

Centre for Epidemiology Versus Arthritis and Versus Arthritis.

Research in context

Evidence before this study

We searched PubMed for articles published in English between Jan 1 1990, and July 13 2023, using the search terms (pain[Title]) AND ((variab*[Title]) OR (fluctuati*[Title]) OR (vary*[Title]) OR (chang*[Title])) AND ((day[Title]) OR (daily[Title])). This search yielded 40 papers. Of these, 17 studies from 16 papers were synthesised. Reasons for exclusion were that papers investigated variability in other symptoms (*n* = 15), did not collect daily data (*n* = 6), did not provide enough information about variability measures (*n* = 1), were not written in English (*n* = 1), and did not study humans (*n* = 1).

Included studies collected data on 1–165 days. Pain variability was explored implicitly in most studies, using multilevel models, dispersion models, linear mixed models, or hierarchical linear models. When pain variability was explicitly described, it was using one measure. This measure was often individual standard deviation or variance (*n* = 5), but also included the range of pain scores (*n* = 1) or a pre-selected percentage increase (*n* = 1).

Pain variability was associated with BMI, site of pain, pain condition, medication, age, race, response to treatment, genotypes, depression, happiness, frustration, and physical activity. Days with changes of pain severity were associated with changes in

fatigue, depressive symptoms, cognitive function, poor sleep, time spent asleep, daily catastrophizing, self-efficacy, negative affect, occupational load, physical activity, and disability. In some studies, there were no statistically significant associations between pain variability and explored variables.

Added value of this study

First, our study expands this previous literature by using multiple measures of pain variability. These measures are suitable for summarising ordinal data and provide a broader picture of the magnitude and temporal nature of pain variability. Second, our previous work assigned weekly trajectories of pain severity data to clusters. The present study reports pain variability within these clusters, supporting a subgroup analysis of pain variability. We report differences between the results of each pain cluster (e.g., that pain trajectories in a no/low pain cluster are more often stable, with no day-to-day fluctuations, than trajectories in other clusters). Third, we identify covariates associated with pain variability among the different clusters. Some predictors of pain change (e.g., pain interference) are identified in all clusters, while other predictors (e.g., lag wind-speed) are identified in only one cluster.

Implications of all the available evidence

Pain variability is an important alternative index to average pain. Digital health technologies allow the opportunity for daily collected data, and these can be analysed to summarise pain variability. Some of the associated covariates with pain variability can be modified. For example, mood can be supported with therapeutic interventions and could therefore be the focus of future work.

Introduction

Digital health technologies (mobile health (mHealth) apps and wearables) can support long-term health conditions[1] by collecting patient-generated health data (PGHD) including exposures and health outcomes. PGHD collected actively (e.g., participants reporting pain severity or mood) and passively (e.g., environmental exposures such as the weather) have many advantages. The data can be collected longitudinally, at high frequency, and in patients' own environments. The data enable the assessment of health states that are often dynamic (i.e., that fluctuate over time, between and within patients). Analysing these data as dynamic processes could provide novel insights.

Chronic pain is common across long-term health conditions, and pain severity is a key outcome measure[2]. Clustering methods are frequently used to identify trajectories of pain severity over time[3]. We have previously used PGHD to identify clusters of weekly pain severity trajectories (unpublished). Using the *k*-medoids algorithm we compared within-week pain severity (but not variability) using the Manhattan distance. These clusters were interpretable in plain language as representing "no/low pain", "mild pain", "moderate pain" and "severe pain". Medoid weekly trajectories of these clusters represented pain severity that did not fluctuate, but spaghetti plots of trajectories highlighted substantial within-week variability (unpublished). Ignoring this variability could mask important pain severity dynamics and relationships with exposures[4,5].

Higher pain variability has been associated with lower quality of life independent of disease severity[6,7], an increased risk of depression[8], but better response to treatment[9]. People living with chronic pain wish to better understand variability in their pain [preprint][10] and report frustration when causes of variability are unknown[11]. The importance of understanding pain variability is ranked higher by patients than by

clinicians and researchers[12]. For these reasons, there is a need to understand pain variability and its associated covariates.

Previous studies have described variability using an individual's standard deviation of their pain severity trajectory. However, this summary measure misses day-to-day fluctuations. To overcome this limitation, using multiple measures of pain variability (including individual standard deviation) have been proposed[13]. Further, data collected through digital health tools offer the opportunity to explore covariates associated with pain variability. The aim of this study was to quantify the day-to-day variability in pain severity and identify covariates associated with that variability.

Methods

Data Source

Data were collected between January 2016 and April 2017 as part of a population-based mHealth study, *Cloudy with a Chance of Pain*. Detailed study methods have been reported elsewhere[14,15]. In brief, people living with long-term pain conditions were recruited through local and national media campaigns[16]. Interested participants downloaded a smartphone app, provided electronic consent, and completed a baseline survey. Included participants had at least one chronic pain condition, were aged ≥17 years, UK residents, and had a mobile smartphone. Then, daily data were collected actively and passively (see below for details) via the app. Data were time- and date-stamped and transferred to secure servers. Overall, 10,483 participants downloaded the app, provided full baseline data, and provided at least one report of pain severity. Ethical approval for *Cloudy with a Chance of Pain* was from the University of Manchester Research Ethics Committee (ref: ethics/15522) and from the NHS IRAS

(ref: 23/NW/0716). Reporting of this secondary analysis follows the STROBE guidelines[17].

Data Items

Baseline data included year of birth, sex (male or female), condition(s) of chronic pain (selected from a list of options, including "other"), and site(s) of chronic pain (selected from a list of options, including "other").

Actively collected PGHD were ten self-reported variables measured using a five-point ordinal scale. The primary outcome was severity of chronic pain. Responses to "How severe was your pain today?", ranged from 1 (no pain) to 5 (very severe pain). Other variables were (with value labels for scores 1 and 5 shown): morning stiffness (no stiffness, very severely stiff), mood (depressed, very happy), fatigue (no fatigue, very severe fatigue), sleep quality (very poor, very good), tiredness on waking (not at all tired, extremely tired), time spent outside (none of the day, all of the day), exercise (no exercise, 30+ minutes strenuous activity), pain interference (not at all, very much), and well-being (very unwell, very well). The scales for mood, sleep quality, time spent outside, exercise, and well-being were inverted for this secondary analysis so that a reduction on the scale represented an improvement and an increase on the scale represented a worsening across all variables.

Passively collected weather data (wind speed, temperature, dewpoint temperature, atmospheric pressure and relative humidity) were collected hourly from the nearest weather station identified using the phone's Global Positioning System (GPS)[18]. Mean daily weather conditions were calculated for each participant based on the hourly data.

Data Processing

Variability in pain severity was examined across a seven-day period, Monday to Sunday. *Available weeks* were defined as weeks where participants had submitted daily pain severity data, Monday to Sunday. There were 21,919 *available weeks* contributed by 2807 participants. For each pair of consecutive days in available weeks, a *between-day change in pain score* was calculated (Figure 1). Similarly, for the same pair of consecutive days, a *between-day change in covariate score* (for each of the nine PGHD variables and the weather variables listed above) were calculated. For those days with missing data, the between-day change in pain or covariate scores was not calculated. Finally, *lagged between-day changes in covariate scores* were calculated for the previous pair of consecutive days. To calculate lagged variables required covariate data to be completed Sunday to Sunday (Figure 1). Changes in each PGHD variable score ranged from –4 to +4. Weather variables were re-scaled with a minimum value of 1 and a maximum value of 5, and weather differences ranged between –4 and +4.

Complete weeks were defined as *available weeks* during which complete covariate and lagged covariate data were available. Participants could contribute up to 64 *complete weeks* (the length of the study). Of the 21,919 *available weeks* contributed by 2807 participants, 13,052 (59·5%) were classified as *complete weeks* contributed by 2070 participants. The analyses presented here were based on these 13,052 complete weeks.

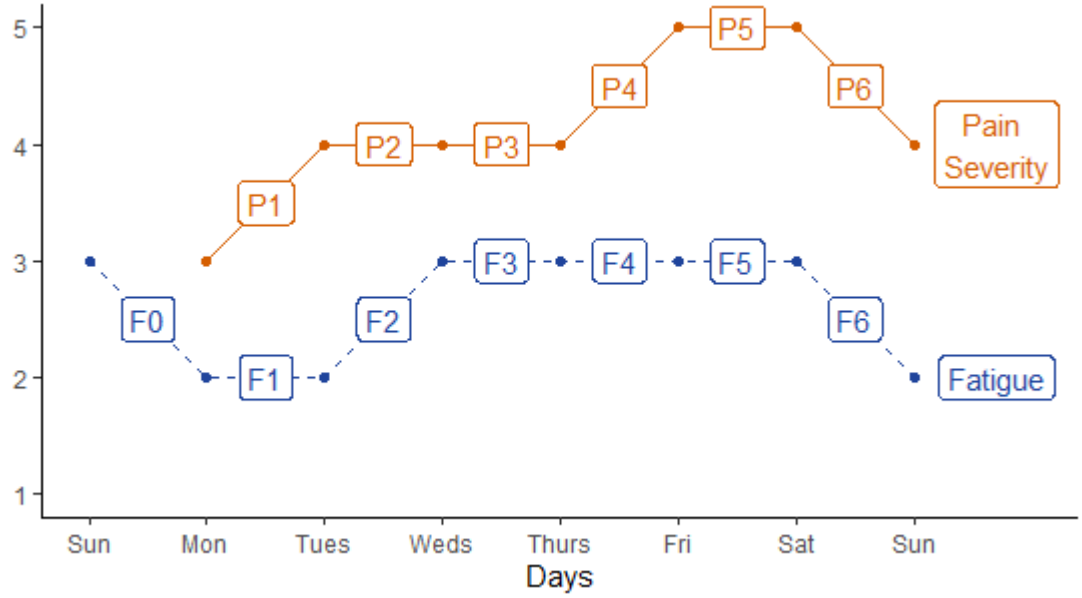


Figure 1: Pain severity and fatigue scores for a hypothetical participant, with between-day changes for a *complete week* (i.e., where a participant has contributed, in this case, pain severity each day Monday to Sunday and fatigue scores each day Sunday to Sunday). Dots indicate daily pain scores (in orange) and daily fatigue scores (in blue). P1 denotes the first between-day change in pain scores, and has a value of +1, P2 denotes the second between-day change in pain scores and has a score of zero, etc. F1 is a concurrent change in fatigue (with value 0) with P1, and F0 is a lagged change in fatigue (with value –1).

For *available* weeks, the mean age of corresponding participants, the proportion of weeks contributed by females, by individuals with each chronic-pain condition, and by individuals with each site of pain were calculated. Sensitivity analyses identified differences in these variables between the *available* weeks that were and were not included as *complete* weeks. To test for a difference in the distributions of ages, a Kolmogorov–Smirnoff test was used. To test for differences in the proportions of weeks contributed by females, by people with each chronic-pain condition, and by people with

each site of pain, chi-squared tests were conducted. When available weeks were incomplete and hence excluded from the analysis, the variable(s) with missing data were recorded.

Statistical Analysis

Quantifying the day-to-day variability in pain severity across a seven-day period

In a cluster analysis of patient-generated pain severity data we reported four clusters with substantial within-cluster variability (unpublished). Figure 2 shows four examples of trajectories within the "moderate pain" cluster which was characterised by an average pain score of three on a one to five scale. Trajectory 1 shows no day-to-day fluctuations, trajectory 2 frequent but moderately sized fluctuations, trajectory 3 infrequent but large fluctuations, and trajectory 4 shows large and frequent fluctuations.

In this study, between-day changes in pain severity were used to quantify variability. A complete week had six between-day changes, taking values between –4 and +4, with 0 being no change. First, weeks were categorised as either a stable-pain week (all between-day pain changes were zero) or a variable-pain week (at least one non-zero between-day pain change). For variable-pain weeks, the following measures of pain variability were calculated:

a) Probability of Acute Change (PAC). The proportion of between-day changes with magnitude ≥1. The PAC for each variable-pain week lies between 0 (a stable trajectory) and 1 (constantly fluctuating pain). The PAC from each variable-pain week within a cluster were used to calculate a mean and standard deviation PAC for that cluster.
b) Autocorrelation coefficient (AC) with lag 1. Takes continuous values between –1 and 1, with values close to –1 representing a 'back-and-forth' trajectory

alternating between being above and below the mean pain severity, and values close to 1 representing trajectories that are similar on consecutive days (e.g., pain above the mean on one day remains above the mean on the following day). The mean and standard deviation AC were reported by cluster.

c) Mean Square of Successive Difference (MSSD). Squares absolute changes in pain severity and then calculates the mean of these values. A higher MSSD can be reported a result of more frequent, or more severe fluctuations. The mean and standard deviation MSSD were reported by cluster.

d) The proportion of changes observed at each value (–4, –3, …, +2, +3, +4) was examined. These proportions were reported as within-cluster distributions, both numerically (on a heatmap) and graphically.

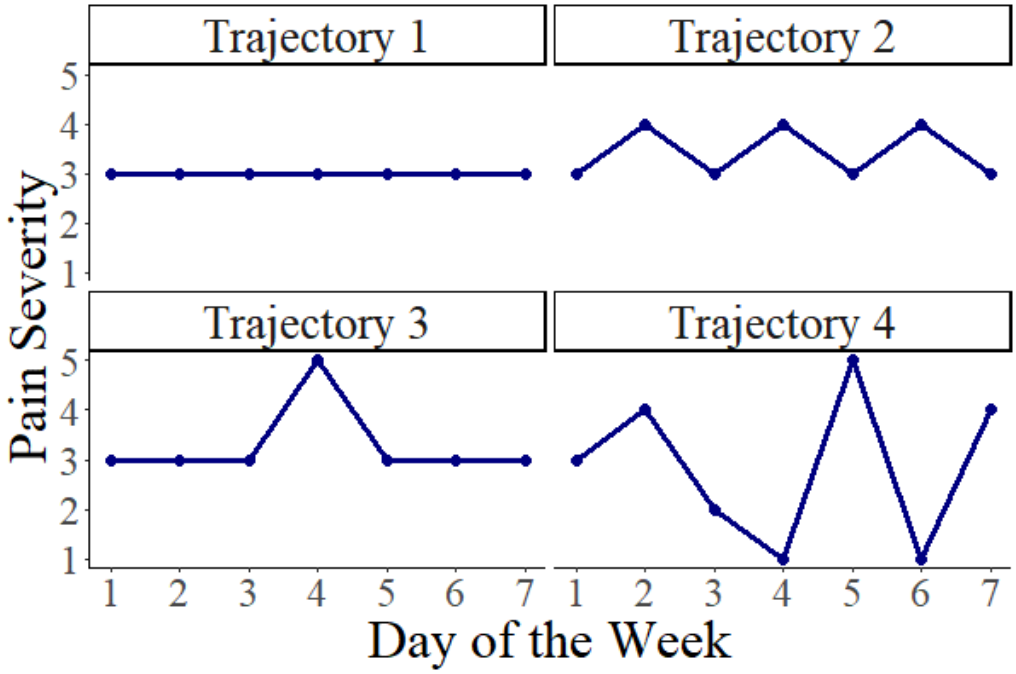


Figure 2. Example trajectory weeks, taken from data previously assigned to Cluster C ("moderate pain"). Trajectory 1 has no variability, trajectory 2 has frequent but small fluctuations, trajectory 3 has infrequent but larger fluctuations, and trajectory 4 has frequent and large fluctuations.

Identifying variables associated with day-to-day variability in pain severity.

The aim of this analysis was to identify variables that changed prior to, or concurrently with, changes in pain severity. Here, the outcome variable was defined as the between-day direction of pain change (i.e., negative, positive or zero). Cumulative probit models were used to model this ordinal outcome[19], using the *ordinalNet* package in *R*[20]. Models were fit separately for each cluster, to avoid masking of associations as per Simpson's Paradox[21], in which opposite associations can be observed at a population-level, than within each subgroup of data.

There were 50 candidate predictors for each of the four models: baseline variables and concurrent and lagged changes in each of the daily covariate variables. Candidate predictors were standardised by subtracting the mean (for continuous variables) or median (for ordinal variables). Due to the large number of predictors, and the possibility of collinearity among predictors, variable selection was conducted for the probit model using the LASSO[22]. LASSO introduces a shrinkage factor to the likelihood function of the probit model. The strength of this shrinkage factor is controlled by the tuning parameter λ. The LASSO algorithm shrinks some coefficients to zero, effectively eliminating them from the model. The LASSO is widely used in variable selection[23] due to its ability to identify groups of important variables to be retained in the model and because it is robust to reasonable collinearity between candidate variables.

To select a model with an appropriate reduction of the variables, 20 values of λ were tested. For each value of λ, a probit model was fitted. Of the 20 candidate models, the model with the smallest Bayesian Information Criterion was selected. This criterion trades off goodness of fit against model complexity and has the benefit of increasing the

complexity penalty for larger datasets such as ours, providing an extra guard against the risk of overfitting to big data.

Model Interpretation

Of the optimal model for each cluster, the predictors that were retained in the model (i.e., whose coefficient was not shrunk to zero) are reported. To quantify the baseline probability of each outcome (pain increase, pain decrease, no pain change), all predictors were set to zero in each optimal model. The baseline probabilities of a pain increase and of a pain decrease are reported. The marginal probability of each retained predictor was calculated by setting all predictors to zero except the predictor of interest. The marginal probability of each predictor was calculated both by increasing and decreasing the predictor by one unit. For each cluster, the four predictors resulting in the largest changes in probability outcomes were reported graphically. Marginal probabilities for all predictors were reported numerically.

Role of the funding source

The funders had no role in the study design, in the collection, analysis or interpretation of data, in the writing of the report, or in the decision to submit for publication.

Results

Data

The proportion of available weeks classified as complete weeks was similar across clusters: Cluster A contained 56.6% (n = 970) complete weeks, Cluster B 59.2% (n = 4885), Cluster C 60.1% (n = 5036), and Cluster D 60.3% (n = 2161). The 2070 participants provided an average of 6.3 complete weeks, and 95% of the participants contributed between 1 and 28 weeks.

Demographic data are provided in Supplementary Table 1. The weeks included in the analyses were significantly more likely to be contributed by participants who were younger, had diagnoses of gout or unspecific arthritis, and reported pain in the stomach/abdomen, hip, knee, or hands, and they were significantly less likely to have diagnoses of fibromyalgia or to report pain in the head, face, mouth/jaws, or neck/shoulder. The missing variables that result in available weeks not being complete are reported in Supplementary Table 2.

Quantifying Variability

Table 1 reports the number and percentage of complete weeks that were stable-pain weeks by cluster. In total, 1474 (11·3%) complete weeks were stable. A higher percentage of weeks were stable in Cluster A (21·9%) when compared to clusters B, C and D. Cluster C (moderate pain) contains the lowest percentage of stable-pain weeks (7·4%).

Table 1: Percentage of stable trajectories		
Cluster	Number (percentage) of complete stable-pain weeks	Description of complete stable-pain weeks
A (low/no pain)	212 (21·9%)	All are constant ‘1’ pain
B (mild pain)	647 (13·2%)	All are constant ‘2’ pain
C (moderate pain)	374 (7·4%)	All are constant ‘3’ pain
D (severe pain)	241 (11·2%)	94 (4·3% of total) are constant ‘4’ pain and 147 (6·8% of total) are constant ‘5’ pain

Among variable-pain weeks, measures of PAC, AC, and MSSD were calculated for each trajectory, with the mean and standard deviation across all trajectories reported in Table 2.

a) PAC. On average, trajectories in Cluster A fluctuated less frequently (0·41) than clusters of more severe pain (e.g., Cluster C: 0·52).
b) AC (with lag 1). In each cluster, the average AC was close to zero. The AC values range from –0·86 to +0·67, with an interquartile range of –0·26 to +0·23.
c) MSSD. Values were greater in clusters of more severe pain severity, indicating more variability in these clusters. Some of this greater variability is captured by the PAC, but it is unclear from these summary measures alone whether one cluster has more severe fluctuations.

Table 2: Measures of variability by cluster

Cluster	PAC among variable trajectories; mean (SD)	AC among variable trajectories; mean (SD)	MSSD among variable trajectories; mean (SD)
A (n=680)	0·409 (0.191)	-0·002 (0·300)	0·63 (0·61)
B (n=3788)	0·474 (0·213)	-0·049 (0·296)	0·67 (0·58)
C (n=4155)	0·517 (0·215)	-0·036 (0·302)	0·77 (0·66)
D (n=1745)	0·500 (0·214)	-0·019 (0·304)	0·80 (0·73)

Among variable-pain weeks, Figure 3 reports the percentage of pain differences observed by cluster (numerically in Figure 3a and graphically in Figure 3b). Complete weeks in Cluster A showed the lowest percentage of changes (59% of consecutive days had no pain change), whereas Cluster C showed the highest percentage of pain changes (48% of consecutive days had no pain change). When pain did change, the change was most likely by an absolute difference of 1 unit, but at least 6% of observed pain differences had a magnitude of ≥2 units. Overall, 0·06% of pain changes were differences of –4 or +4, and were more likely in Clusters A (low/no pain) and D (severe pain), likely due to pain severity in these clusters being at extreme ends of the scale, and therefore permitting greater movement.

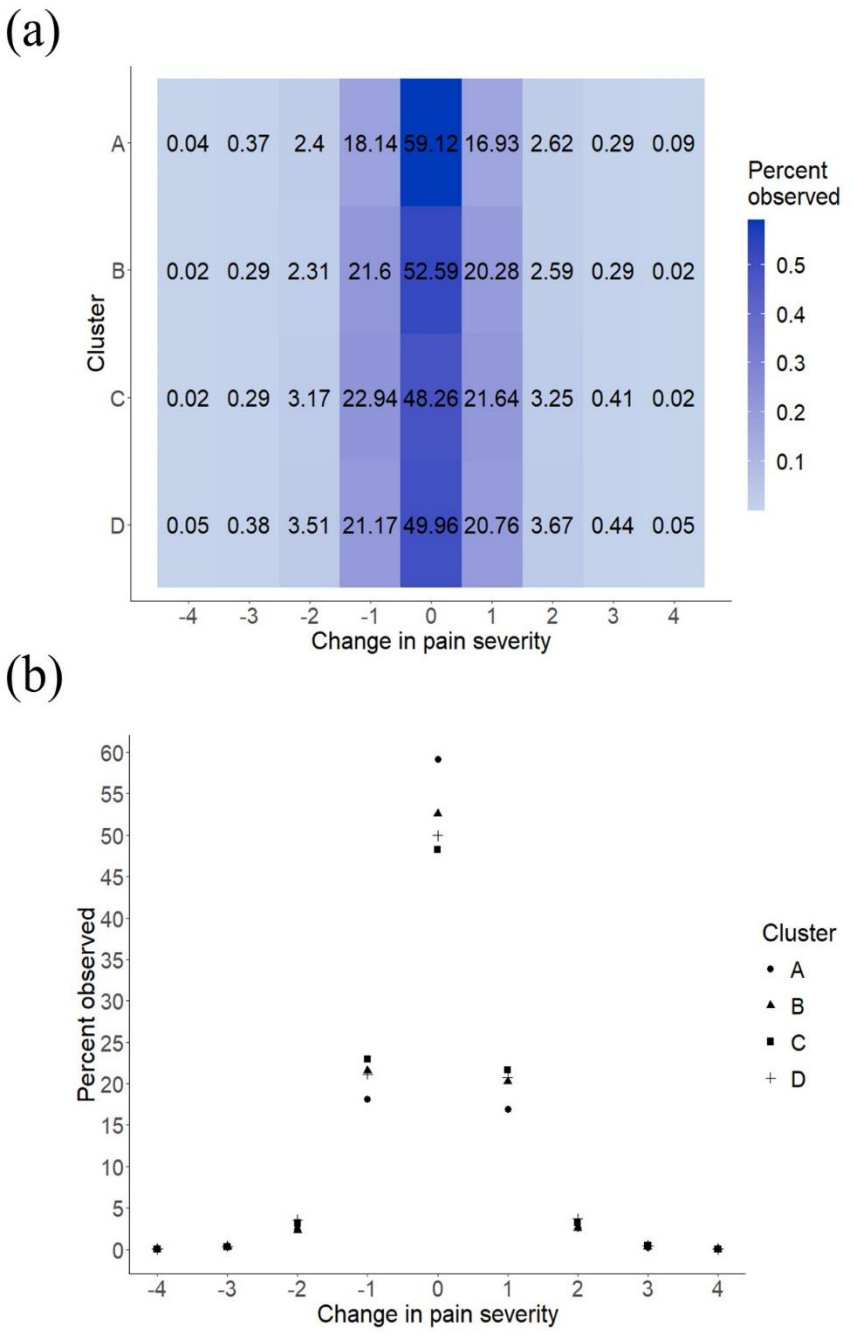


Figure 3: Distribution of pain changes observed within variable-pain weeks. Figure (a) shows the percentage of each value numerically, Figure (b) shows the same data graphically.

Identifying variables associated with pain variability

For each optimal model identified, and given no change in covariates, the baseline probabilities of experiencing an increase or decrease in pain severity are reported in Table S3 and visualised by horizontal lines in Figure 4. With no changes in predictor variables, day-to-day pain changes had a probability of improvement of 0·128 in Cluster A, 0·172 in Cluster B, 0·202 in Cluster C, and 0·182 in Cluster D, and a probability of worsening of 0·120 in Cluster A, 0·163 in Cluster B, 0·192 in Cluster C, and 0·180 in Cluster D.

After LASSO variable selection with different levels of shrinkage (λ), fatigue, mood, stiffness upon waking, sleep quality, pain interference and patient well-being were retained in all models, indicating associations between concurrent changes in these variables and in pain severity across all clusters of pain severity (Table 3).

Table 3: Variables associated with changes in pain

Cluster	A	B	C	D
Fatigue	■	■	■	■
Mood	■	■	■	■
Exercise		■	■	■
Time spent outside			■	■
Morning stiffness	■	■	■	■
Lag morning stiffness		■		
Sleep quality	■	■	■	■
Pain inteference	■	■	■	■
Lag pain inteference		■	■	
Well-being	■	■	■	■
Lag wind speed			■	

Variables retained by models following shrinkage through LASSO are shown by the blue colouring. Changes are concurrent with pain change unless otherwise stated.

Supplementary Table 3 shows the marginal probability of pain worsening or improving given exposure data with non-zero coefficients. Figure 4 shows the four variables with the largest changes in marginal probabilities: changes in pain interference, morning stiffness, well-being and fatigue. A one-unit change in pain interference had the greatest impact on the probability of a change in pain severity, held true across all clusters. Increases in pain interference were associated with increases in the probability of day-to-day pain worsening to 0·395 in Cluster A (up by 0·275 from baseline), 0·340 in Cluster B (up by 0·177 from baseline), 0·343 in Cluster C (up by 0·151 from baseline), and 0·301 in Cluster D (up by 0·121 from baseline). Increases in pain interference were also associated with decreases in the probability of day-to-day pain improvement to 0·020 in Cluster A (down by 0·108 from baseline), 0·065 in Cluster B (down by 0·107 from baseline), 0·097 in Cluster C (down by 0·105 from baseline), and 0·097 in Cluster D (down by 0·085 from baseline).

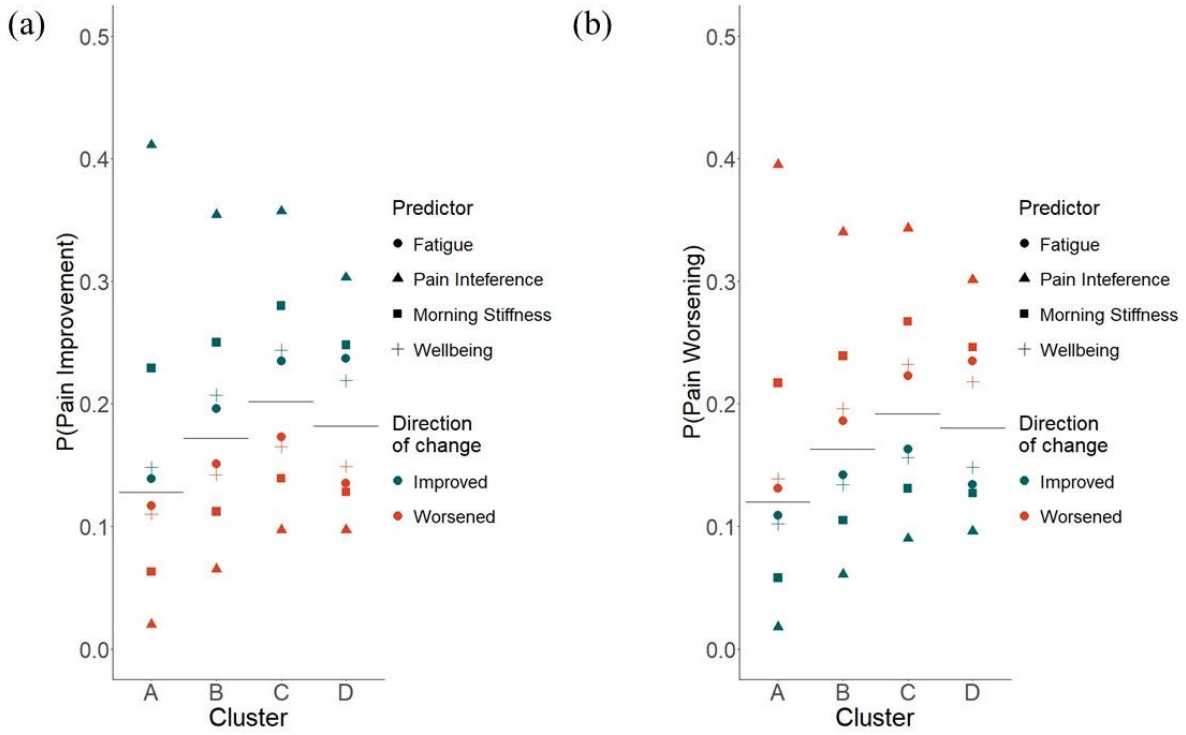


Figure 4: (a) Baseline (no change in PGHD and weather variables) probability of pain severity improving shown by horizontal line for each cluster. (b) Baseline probability of pain severity worsening shown by horizontal line for each cluster. Points represent probability of pain severity improving given an improvement of one-unit (green) or worsening of one-unit (red) of fatigue, pain interference, morning stiffness, and wellbeing.

Discussion

Summary

Daily data from an mHealth study were used to quantify variability in pain severity across seven-day periods. About 10% of weeks were stable (with no day-to-day changes in pain severity). Within-clusters, variable-pain weeks had no pain changes in 50–60% of consecutive days. Non-zero changes were most likely to be one-unit (of a maximum 4) in magnitude, with 6% being ≥2 units.

Across all clusters, the average autocorrelation was close to zero, but with a large range and interquartile range, therefore no population-level conclusion could be drawn from this measure. The PAC indicated pain changes on around half of observed days. The MSSD taken alongside the distribution of pain changes indicated that more severe pain clusters (Cluster C and Cluster D) exhibited more severe fluctuations. However, the more extreme clusters (Clusters A and D) observed the highest proportion of extreme changes (–4 or +4), likely due to the opportunity for greater movement between pain severity scores in these clusters. The results suggest that if pain severity is low over a week (Cluster A), then it is more likely to be stable, whereas weeks with moderate or severe pain (Cluster C and Cluster D) are more likely to fluctuate. Across all clusters changes in pain interference, well-being, morning stiffness, fatigue, sleep quality, and mood were associated with changes in pain severity. For other predictors there was evidence of cluster-specific associations (e.g. pain changes were associated with lag morning stiffness in Cluster B only and with lag wind speed in Cluster C only).

Context

Previous work has measured pain variability by using a single statistic, most commonly the within-trajectory standard deviation. This statistic summarises the magnitude of distance from the mean but has no temporal features and so does not measure day-to-day fluctuations. Mun et al.[13] described three other measures of pain variability (PAC, AC, and MSSD), and we have explored these in this study. Each of these measures has drawbacks for understanding pain variability. Autocorrelation summarises temporal characteristics, but not the magnitude, of changes. MSSD is difficult to interpret for participants and clinicians. PAC requires a subjective cut-off point. Summarising all of these, as well as distributions of pain variability, has provided a broader description of pain variability than could be afforded by any single measure.

Similar to our study, Wesolowicz et al.[24] reported that low pain severity was associated with low pain variability. While associations between pain variability and exposure data (mental and physical health and functioning) have been previously reported[25], some studies[26,27] found no predictors of pain variability. Our study reported that in all clusters of pain severity, pain changes were associated with changes in pain interference, morning stiffness, patient well-being, and fatigue.

Strengths and Limitations

This study has several strengths. We described pain variability using multiple measures. Avoiding the use of a single summary statistic provided a more holistic description of pain variability. We used previously identified clusters to identify differences in pain variability among these clusters (in this case, based on the severity of pain). Exploring variability at a group-level avoided masking of population-level associations and identified cluster-specific associations. Finally, we tested associations of pain change with both concurrent and lagged variables, including passively collected data.

This study also had several limitations. First, only complete data were used. Data may not be missing at random and there may be differences between complete weeks, and those that were available but with missing data. For example, people who did not provide complete weeks may have had stable pain and did not perceive a benefit to contributing the same data each day. Second, there were differences between the demographic data of available weeks and those that were complete. This may mean that the observed associations are particularly relevant for the contributors to this analysis, but not generalisable to the wider study population. Third, the weeks used in this study were defined to be between Monday and Sunday. Previous analyses in identifying clusters of pain severity found no difference between weeks described in this way, and weeks defined across other spans (e.g., Tuesday–Monday), but this may not remain true

for pain variability. Fourth, data collection was once daily. There is evidence that pain varies within-days,[28] and this has not been captured by the analysis in this study. Fifth, the LASSO method is generally robust for improving model performance in the presence of collinear predictors but the increased suitability of selected covariates over deselected collinear covariates is unclear. Sixth, the LASSO method and Bayesian Information Criterion were used to limit overfitting but no formal sample size calculation was performed and there hence remains the potential for overfitting due to the large number of candidate predictors relative to the number of events of pain worsening and of pain improvement.

Future directions/conclusions

Previous work has reported that participants want to know about future fluctuations in pain severity [preprint][10]. In this paper, we have shown that daily collected PGHD can be used to summarise pain variability and associated predictors of that variability. The methods for summarising pain variability in this study are not specific to chronic pain and could be used in other fields (e.g., mental health) to explore variability in daily symptoms. Future work can use the identified associations to explore the feasibility of forecasting future fluctuations in pain severity. Some of the identified associations are modifiable (e.g., mood can be improved with psychological support), and could be the focus of future work to reduce pain variability.

Contributions:

All authors contributed to the conception of the study. CLL and BBY designed the statistical methods. CLL accessed and verified the data analysed in this manuscript. All authors had access to this data. CLL conducted statistical analysis of the data and wrote the first draft of the manuscript. All authors edited the manuscript and agreed to submit for publication.

Data sharing:

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request. Data management and analyses were performed in R 4.1.2. Code is available on reasonable request.

Declaration of interests:

WGD has received consultancy fees from Google, and DMS has received consultancy fees from Palta, both unrelated to this work. All other authors have no conflicts of interest to declare.

Acknowledgements:

This study was supported by the Centre for Epidemiology Versus Arthritis (grant number 21755). Cloudy with a Chance of Pain was funded by Versus Arthritis (grant reference 21225), with additional support from the Centre for Epidemiology (grants 21755 and 20380). TH receives funding from the Royal Society (grant number INF/R2/180067) and the Alan Turing Institute for Data Science and Artificial Intelligence. The funders of the study had no role in study design, data interpretation, writing of the manuscript, or the decision to submit for publication.

References

1. Solomon DH, Rudin RS. Digital health technologies: opportunities and challenges in rheumatology. *Nat Rev Rheumatol*. 2020;**16**(9):525-535. doi:10.1038/s41584-020-0461-x

2. McQueenie R, Jani BD, Siebert S, et al. Prevalence of chronic pain in LTCs and multimorbidity: A cross-sectional study using UK Biobank. *J Multimorb Comorbidity*. 2021;**11**:263355652110058. doi:10.1177/26335565211005870

3. Kongsted A, Kent P, Axen I, Downie AS, Dunn KM. What have we learned from ten years of trajectory research in low back pain? *BMC Musculoskelet Disord*. 2016;**17**:220. doi:10.1186/s12891-016-1071-2

4. Fillingim RB, Loeser JD, Baron R, Edwards RR. Assessment of Chronic Pain: Domains, Methods, and Mechanisms. *J Pain*. 2016;**17**(9 Suppl):T10-20. doi:10.1016/j.jpain.2015.08.010

5. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*. 2005;**113**(1):9-19. doi:10.1016/j.pain.2004.09.012

6. Tupper SM, Rosenberg AM, Pahwa P, Stinson JN. Pain Intensity Variability and Its Relationship With Quality of Life in Youths With Juvenile Idiopathic Arthritis. *Arthritis Care Res*. 2013;**65**(4):563-570. doi:10.1002/acr.21850

7. Zakoscielna KM, Parmelee PA. Pain variability and its predictors in older adults: depression, cognition, functional status, health, and pain. *J Aging Health*. 2013;**25**(8):1329-1339. doi:10.1177/0898264313504457

8. Schneider S, Junghaenel DU, Keefe FJ, Schwartz JE, Stone AA, Broderick JE. Individual differences in the day-to-day variability of pain, fatigue, and well-being in patients with rheumatic disease: associations with psychological variables. *Pain*. 2012;**153**(4):813-822. doi:10.1016/j.pain.2012.01.001

9. Bonin Pinto C, Barroso J, Schinitzer TJ. Characterization and implications of daily pain variability and response to treatment in osteoarthritis clinical trials. *Osteoarthritis Cartilage*. 2021;**29**:S277. doi:10.1016/j.joca.2021.02.364

10. Little CL, Druce KL, Dixon WG, Schultz DM, House T, McBeth J. *What Do People Living with Chronic Pain Want from a Pain Forecast? A Research Prioritisation Study*. Epidemiology; 2023. doi:10.1101/2023.04.24.23289032

11. Pagé MG, Dassieu L, Develay E, et al. The Stressful Characteristics of Pain That Drive You NUTS: A Qualitative Exploration of a Stress Model to Understand the Chronic Pain Experience. *Pain Med Malden Mass*. 2021;**22**(5):1095-1108. doi:10.1093/pm/pnaa370

12. Stone AA, Broderick JE, Goldman RE, et al. I. Indices of Pain Intensity Derived From Ecological Momentary Assessments: Rationale and Stakeholder Preferences. *J Pain*. 2021;**22**(4):359-370. doi:10.1016/j.jpain.2020.08.003

13. Mun CJ, Suk HW, Davis MC, et al. Investigating intraindividual pain variability: methods, applications, issues, and directions. *Pain*. 2019;**160**(11):2415-2429. doi:10.1097/j.pain.0000000000001626

14. Dixon WG, Beukenhorst AL, Yimer BB, et al. How the weather affects the pain of citizen scientists using a smartphone app. *NPJ Digit Med*. 2019;**2**:105. doi:10.1038/s41746-019-0180-3

15. Schultz DM, Beukenhorst AL, Yimer BB, et al. Weather Patterns Associated with Pain in Chronic-Pain Sufferers. *Bull Am Meteorol Soc*. 2020;**101**(5):E555-E566. doi:10.1175/BAMS-D-19-0265.1

16. Druce KL, Dixon WG, McBeth J. Maximizing Engagement in Mobile Health Studies: Lessons Learned and Future Directions. *Rheum Dis Clin North Am*. 2019;**45**(2):159-172. doi:10.1016/j.rdc.2019.01.004

17. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Bull World Health Organ*. 2007;**85**(11):867-872. doi:10.2471/blt.07.045120

18. Beukenhorst AL, Sergeant JC, Schultz DM, McBeth J, Yimer BB, Dixon WG. Understanding the Predictors of Missing Location Data to Inform Smartphone Study Design: Observational Study. *JMIR MHealth UHealth*. 2021;**9**(11):e28857. doi:10.2196/28857

19. Agresti A. *Categorical Data Analysis*. 3rd ed. Wiley; 2013.

20. Wurm MJ, Rathouz PJ, Hanlon BM. Regularized Ordinal Regression and the **ordinalNet** *R* Package. *J Stat Softw*. 2021;**99**(6). doi:10.18637/jss.v099.i06

21. Simpson EH. The Interpretation of Interaction in Contingency Tables. *J R Stat Soc Ser B Methodol*. 1951;13(2):238-241. doi:10.1111/j.2517-6161.1951.tb00088.x

22. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J R Stat Soc Ser B Methodol*. 1996;**58**(1):267-288.

23. Ranstam J, Cook JA. LASSO regression. *Br J Surg*. 2018;**105**(10):1348-1348. doi:10.1002/bjs.10895

24. Wesolowicz DM, Bishop MD, Robinson ME. An Examination of Day-to-Day and Intraindividual Pain Variability in Low Back Pain. *Pain Med Malden Mass*. 2021;**22**(10):2263-2275. doi:10.1093/pm/pnab119

25. Schneider S, Junghaenel DU, Broderick JE, Ono M, May M, Stone AA. II. Indices of Pain Intensity Derived From Ecological Momentary Assessments and Their Relationships With Patient Functioning: An Individual Patient Data Meta-analysis. *J Pain*. 2021;**22**(4):371-385. doi:10.1016/j.jpain.2020.10.002

26. Pagé MG, Gauvin L, Sylvestre MP, Nitulescu R, Dyachenko A, Choinière M. An Ecological Momentary Assessment Study of Pain Intensity Variability: Ascertaining Extent, Predictors, and Associations With Quality of Life, Interference and Health Care Utilization Among Individuals Living With Chronic Low Back Pain. *J Pain*. 2022;23(7):1151-1166. doi:10.1016/j.jpain.2022.01.001

27. Schneider S, Junghaenel DU, Keefe FJ, Schwartz JE, Stone AA, Broderick JE. Individual differences in the day-to-day variability of pain, fatigue, and well-being in patients with rheumatic disease: associations with psychological variables. *Pain*. 2012;**153**(4):813-822. doi:10.1016/j.pain.2012.01.001

28. Allen KD, Coffman CJ, Golightly YM, Stechuchak KM, Keefe FJ. Daily pain variations among patients with hand, hip, and knee osteoarthritis. *Osteoarthritis Cartilage*. 2009;**17**(10):1275-1282. doi:10.1016/j.joca.2009.03.021

6. Manuscript 4: Forecasting pain severity using daily patient-generated health data collected using a smartphone app

Chapter 2 reported that people with chronic pain wanted a forecast that could predict pain flares. Movement to worsened pain clusters could represent these pain flares. Therefore, forecasting transitions between clusters are explored in this chapter. Variability measures explored in Chapter 5 are among the candidate predictors for these transitions.

I led discussions with the coauthors to conceptualize the work presented in this chapter. I primarily developed the methodology with support in technical aspects from coauthor TH. I developed the code used for the analysis and conducted the statistical analysis. I wrote the original draft of the manuscript. All coauthors reviewed and edited the manuscript.

The manuscript presented in this chapter has been prepared for submission to *npj Digital Medicine.*

Forecasting pain severity using daily patient-generated health data collected using a smartphone app

Target Journal: npj Digital Medicine

Notes to authors:

- Main text (excluding Methods) is typically no more than 4000-4500 words
- Articles have up to 10 display items (Figures/Tables)
- Abstracts are not usually more than 150 words
- Introduction-Results-Discussion-Methods-Data Availability

Authors:

Claire L Little, MA

ORCID: 0000-0003-2602-0491

Centre for Epidemiology Versus Arthritis,
University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
claire.little@postgrad.manchester.ac.uk

Prof Thomas House, PhD
ORCID: 0000-0001-5835-8062
Department of Mathematics
University of Manchester, Manchester, United Kingdom

Prof David M Schultz, PhD
ORCID: 0000-0003-1558-6975
Department of Earth and Environmental Sciences,
Centre for Atmospheric Science,
University of Manchester, Manchester, United Kingdom
Centre for Crisis Studies and Mitigation
University of Manchester, Manchester, United Kingdom

Prof William G Dixon, MRCP, PhD
ORCID: 0000-0001-5881-4857
Centre for Epidemiology Versus Arthritis,
University of Manchester, Manchester, United Kingdom
NIHR Manchester Musculoskeletal Biomedical Research Unit,
Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Prof John McBeth, PhD
ORCID: 0000-0001-7047-2183
Centre for Epidemiology Versus Arthritis,
University of Manchester, Manchester, United Kingdom
NIHR Manchester Musculoskeletal Biomedical Research Unit,
Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Abstract

Pain is a common symptom of many long-term health conditions. Digital health technologies enable longitudinal collection of pain symptom data, to monitor and support people living with chronic pain. Previous work used such data to describe four clusters of weekly pain trajectories between which there remained substantial between-cluster movement across consecutive weeks. Here we present a multinomial model for forecasting such between-cluster movement. Our approach avoids false discovery through use of an elastic net penalty to identify predictors from a list of candidates. Results show improvement in predictive power over a benchmark model but with movement between clusters having a large element of random chance. Different predictors were identified for each origin cluster, highlighting the importance of performing predictive analysis on a subgroup level.

Introduction

Pain is a common symptom of many musculoskeletal and rheumatic diseases. Conditions with pain as a primary symptom are among the global leading causes of years lived with disability[1]. Higher pain severity is associated with lower quality of life[2,3], worse mental health outcomes[4,5], and loss of earnings due to required changes in work life[6]. Living with chronic pain permeates daily living and requires constant self-management[7,8].

Digital health technologies, which incorporates the use of mobile applications (apps) and wearable technology[9], offer an opportunity to support people living with chronic pain. Mobile health (mHealth) enables the collection of patient-generated health data[9] including symptoms such as pain. These data can be captured longitudinally, granularly, and in participants' own environments. Patient-generated health data can inform patient–clinician conversations, empower patients, and support pain self-management[10]. One promising avenue of research is to forecast future health states. Patient-generated health data have been used to forecast one-day ahead symptom scores in hayfever[11], one-day ahead exacerbation events in eczema[12], and three-day ahead events in chronic obstructive pulmonary disease[13]. This field is emerging and there is no consensus on the most appropriate method(s) for forecasting symptom outcomes using patient-generated health data. However, the results of previous forecasting systems have had low predictive power, especially among models using individual data with few data points[11] or small groups of candidate predictors[12]. Therefore, candidate models should use groups of participant data and should investigate increasingly large groups of predictors.

People with chronic pain have expressed the need for forecasting their pain symptoms, especially for predicting flares in pain severity[14]. We previously used data from the mHealth study *Cloudy with a Chance of Pain*[15] to describe four clusters of weekly pain

trajectories defined by the severity of pain (no/low pain, mild pain, moderate pain, severe pain; paper in preparation, CLL, DMS, TH, WGD, JMcB). Up to 70% of consecutive weekly pain trajectories provided by the same participant remained in the same cluster, and one third transitioned to a new cluster. One approach to forecasting pain would be to use these week-to-week transitions between common patterns of pain severity, including those that lead to worse pain, to predict one week ahead pain severity.

The aim of this study was to develop a model to forecast one week ahead pain severity and identify the predictors of changes in pain severity. The specific objectives were to (1) optimise tuning parameters for predictor selection among candidate models, (2) select the optimal model from those tested, (3) describe predictors in the optimal model associated with movement between clusters.

Results

This analysis was conducted using data from the Cloudy with a Chance of Pain study[15]. At baseline 10,584 participants with chronic pain reported age, sex, diagnosed pain condition(s), and number of bodily sites in which they were currently experiencing pain[15]. On each day for up to 15 months participants reported the severity of their pain, as well as pain interference, sleep quality, fatigue, waking up feeling tired, mood, morning stiffness, wellbeing, levels of physical activity, and time spent outside. GPS-derived location-based weather parameters (wind speed, temperature, dew point, relative humidity, and pressure) were collected concurrently.

The unit of analysis was clusters of trajectories of weekly pain severity scores. We identified the predictors of movement from the starting cluster of weekly pain severity trajectories (here called the “origin cluster”) to the subsequent next-week cluster (the

"destination cluster"). Trajectories were assigned to origin clusters and included for analysis if there was a corresponding trajectory in a destination cluster and sufficient data to examine all candidate predictors (inclusion criteria described in detail in "Methods"). These criteria resulted in the inclusion of 11,469 weeks contributed by 1685 participants. Trajectories meeting the inclusion criteria were classified as no/low pain (cluster A, n = 908), mild pain (cluster B, n = 4164), moderate pain (cluster C, n = 4323), and severe pain (cluster D, n = 1937). Of the included trajectories, the percentage contributed by females was 82.6%, the average age of corresponding participants was 52.9 (standard deviation (SD) = 12.3), the average number of pain conditions was 1.7 (SD = 0.8), and the average number of pain sites was 2.7 (SD = 0.6).

Predicting movement from origin to destination clusters

We modelled the movement from origin to destination clusters. Separate models were developed for each origin cluster, with the destination cluster being the outcome of interest. Movement between clusters A (no/low pain) and D (severe pain) were rare (0.3% from A to D, and 0.06% from D to A). For that reason, for origin Cluster A we modelled movement to destination cluster A, B or C/D. For origin Cluster D we modelled movement to cluster A/B, C or D.

The role of 55 candidate predictors across five categories were examined (sociodemographic variables, n = 4; pain characteristics, n = 4; weather in the origin cluster week, n = 10; self-reported symptoms n = 27; and weather in the destination cluster week n = 10). These candidate predictors included measures of central tendency (mean, median) and measures of spread (standard deviation (SD), probability of acute change (PAC), mean square of successive difference (MSSD)).

Using these candidate predictors, six models of increasing complexity were constructed. First a cluster-only model was constructed using the 'last value carried forward' (i.e., the origin and destination clusters were the same). Then five multinomial models (Models 1–5) with an elastic net penalty sequentially added the sociodemographic, pain characteristics, weather (origin cluster), self-reported symptoms, and weather (destination cluster) predictors.

Objective 1: Optimizing tuning parameters for predictor selection among candidate models.

To achieve model parsimony, we fitted the multinomial models 1–5 with an elastic net penalty term. The penalty includes a proportion of a LASSO penalty which acts as feature selection and a ridge regression penalty which shrinks coefficients toward zero. The hyperparameters (numbers used to specify the exact structure of the model penalty and optimised for predictive performance) of these penalty terms are reported in Supplementary Table 1. Here, α is a measure of the proportion of the elastic net that is a LASSO penalty (compared to a ridge regression penalty) with values ranging from 0 (pure ridge regression penalty term) to 1 (representing a pure LASSO penalty term). Further, λ impacts the size of the overall elastic net penalty term.

For each origin cluster, $\alpha = 0$ for Model 1 (i.e., included sociodemographic predictors). This represents a pure ridge regression penalty term. In 9 of the remaining 16 candidate models, $0 < \alpha < 1$ indicating that a balance between LASSO and ridge regression had been reached. None of the candidate models had $\alpha = 1$, meaning that the LASSO penalty wasn't exclusively used in any model. The hyperparameters for each of the optimal models are presented in Objective 2.

Objective 2: Selecting the optimal model to predict movement between origin and destination clusters.

The misclassification rates (i.e., the percentage of destination cluster trajectories that were incorrectly classified) for all models are reported in Table 1. For the cluster-only model which predicted that the origin and the destination clusters would be the same, the misclassification rates were 36.1% for cluster A, 32.5% for cluster B, 37.5% for cluster C, and 30.0% for cluster D. These figures show that remaining in cluster was the most common state, with between 62 to 70% of trajectories correctly predicted. However, between 30-40% of trajectories had moved to a different cluster and were incorrectly classified. The inclusion of predictor variables improved model fit in clusters A, C and D, evidenced by a reduction in misclassification rate of 0.5–2.8%. The largest reductions in misclassification rate were observed for predicting movement out of origin Cluster A (no/low pain) and D (severe pain). The smallest improvements were observed for predicting movement out of origin Cluster B with all values <0.1%.

Table 1: Misclassification rates by cluster by model

Model	Description	Cluster A	Cluster B	Cluster C	Cluster D
Cluster-only model	No predictor variables	36.1	32.5	37.5	30.0
Model 1	Sociodemographic	36.1	32.5	37.5	30.0
Model 2	Model 1 + pain characteristics	35.6	32.5	37.1	27.9
Model 3	Model 2 + weather in origin cluster week	35.3	32.5	37.1	27.4
Model 4	Model 3 + self-reported symptoms	34.7*	32.5	36.9*	27.1*
Model 5	Model 4 + weather in destination cluster week	34.9	32.5*	37.0	27.2

For the baseline model and each multinomial model, the misclassification rate (percentage of trajectories predicted to a different destination cluster than observed) is reported. All values are %. * indicates the smallest within-cluster misclassification rate.

Cluster only model – predicted that the origin and the destination clusters would be the same

Model 1 – included sex, number of pain conditions, number of pain sites, and age as predictors.

Model 2 – included Model 1 predictors plus pain severity in origin week: maximum, minimum, PAC, and MSSD.

Model 3 – included Model 2 predictors plus wind-speed, temperature, dewpoint temperature, pressure, relative humidity in origin week: mean, SD.

Model 4 – included Model 3 predictors plus fatigue, mood, morning stiffness, pain interference, well-being, physical activity, sleep quality, time spent outside, waking up tired in origin week: median, PAC, MSSD.

Model 5 – included Model 4 predictors plus wind-speed, temperature, dewpoint temperature, pressure, relative humidity in destination week: mean, SD.

The optimal model for each origin cluster (i.e., the models with the lowest misclassification error) were Model 4 for clusters A, C and D, and Model 5 for cluster B. For these models, confusion matrices (counts for each combination of observed and predicted number of trajectories that would move to each destination cluster) are reported in Table 2. Counts on the diagonal of the matrix (Tables 2 a, b, c and d, highlighted) are trajectories that have been accurately forecast. Overall, the models correctly predicted the destination cluster 67.7% of the time when the origin cluster was A, 68.0% for origin cluster B, 63.5% for origin cluster C, and 75.3% for origin cluster D. When the origin and destination clusters were the same (i.e., there was no observed movement between clusters), 94.9% of destination trajectories in Cluster A, 97.5% in Cluster B, 96.0% in Cluster C, and 90.0% in Cluster D were correctly assigned by the optimal model. When the origin and destination clusters were different (i.e., there was observed movement between clusters), the optimal model correctly assigned 22.6% of destination trajectories in Cluster A, 6.6% in Cluster B, 9.4% in Cluster C, and 41.2% in Cluster D.

Table 2: Confusion matrices for trajectories moving from (a) Cluster A, (b) Cluster B, (c) Cluster C, and (d) Cluster D.

a) Origin cluster A, *n* = 908 trajectories					
		Observed destination cluster			
		A	B	C/D	**Total**
Predicted destination cluster	A	541	230	18	789
	B	39	74	6	119
	C/D	0	0	0	0
	Total	570	304	24	908
Confusion matrix for trajectories with origin cluster A. Each cell shows a count of the trajectories in each combination of predicted and observed destination clusters. Grey squares show counts when the predicted cluster matched the observed cluster.					

b) Origin cluster B, *n* = 4,164 trajectories						
		Observed destination cluster				
		A	B	C	D	**Total**
Predicted destination cluster	A	3	3	0	0	6
	B	314	2,741	893	44	3,992
	C	2	68	86	10	166
	D	0	0	0	0	0
	Total	319	2,812	979	54	4,164
Confusion matrix for trajectories with origin cluster B. Each cell shows a count of the trajectories in each combination of predicted and observed destination clusters. Grey squares show counts when the predicted cluster matched the observed cluster.						

c) Origin cluster C, *n* = 4,323 trajectories						
		Observed destination cluster				
		A	B	C	D	**Total**
Predicted destination cluster	A	0	0	0	0	0
	B	26	144	100	11	281
	C	13	908	2,593	512	4,026
	D	0	1	7	8	16
	Total	39	1,053	2,700	531	4,323
Confusion matrix for trajectories with origin cluster C. Each cell shows a count of the trajectories in each combination of predicted and observed destination clusters. Grey squares show counts when the predicted cluster matched the observed cluster.						

d) Origin cluster D, *n* = 1,563 trajectories					
d)		**Observed destination cluster**			
		A/B	C	D	**Total**
Predicted destination cluster	A/B	12	2	4	18
	C	19	228	132	379
	D	18	303	1,219	1,540
	Total	49	533	1,355	1,937
Confusion matrix for trajectories with origin cluster D. Each cell shows a count of the trajectories in each combination of predicted and observed destination clusters. Grey squares show counts when the predicted cluster matched the observed cluster.					

Objective 3: Predictors in the optimal model associated with movement between clusters

For each optimal model, the corresponding values of α and λ were used to determine coefficients of the predictors from origin to destination cluster (Table S2). For origin cluster A, Model 4 was the optimal model (Table 1), it used pure ridge regression ($\alpha = 0$, Table S1) and all candidate predictors from Model 4 (i.e., excluding weather in the destination week) were included in the model (Table S2). Therefore, no coefficients were set to zero, but some were shrunk toward zero. For origin cluster B, Model 5 was the optimal model, and in clusters C and D, models 4 were the optimal model, they all used a combination of LASSO and ridge regression penalties ($0 < \alpha < 1$, Table S1). The coefficients of some covariates were set to zero, effectively removing these covariates from the model. Of the optimal models for movement between origin and destination clusters, the value of λ was the largest for origin cluster A. This indicates that this model experienced the greatest amount of shrinkage among the coefficients (possibly because no feature selection occurred for this origin cluster).

Table 3 shows, separately for each origin cluster, the three predictors associated with the largest change to the probability of movement to the destination cluster. The

direction of association is also shown. Positive associations indicate that an increase in the predictor is associated with an increase in the probability of the corresponding cluster movement. Negative associations indicate that a decrease in the predictor is associated with an increase in the probability of the corresponding cluster movement. For example, among trajectories in origin cluster A, a decrease in median fatigue is associated with an increased probability of remaining in cluster A, while an increase in median fatigue is associated with an increased probability of movement to clusters C/D.

Each movement has different combinations of important covariates for prediction but there are some similarities: measures of pain severity (n = 9), pain interference (n = 6), morning stiffness (n = 4), fatigue (n = 5), dewpoint temperature (n = 3), and the number of pain conditions (n = 3) were all within the key predictors of at least three movements. At least two weather covariates were key predictors of movement from origin Cluster D to each other cluster, but weather covariates were not identified as key predictors from other origin clusters.

Table 3: Largest coefficients for predicting movement between clusters					
		Destination cluster			
		A	**B**	**C**	**D**
Origin cluster	**A**	1. Morning stiffness: median (-) 2. Fatigue: median (-) 3. Waking up tired: median (-)	1. Morning stiffness: median (+) 2. Pain interference: median (-) 3. Number of pain conditions (-)	1. Pain interference: median (+) 2. Fatigue: median (+) 3. Fatigue: MSSD (+)	
	B	1. Pain severity: minimum (-) 2. Pain interference: median (-) 3. Fatigue: median (-)	1. Pain severity: maximum (-) 2. Sex (-) 3. Pain severity: PAC (-)	1. Pain severity: minimum (+) 2. Pain interference: median (+) 3. Number of pain conditions (+)	1. Fatigue: median (+) 2. Pain severity: maximum (+) 3. Pain interference: MSSD (+)
	C	1. Pain severity: minimum (-) 2. Fatigue: median (-) 3. Number of pain sites (-)	1. Pain severity: minimum (-) 2. Pain severity: maximum (-) 3. Morning stiffness: median (-)	1. Pain severity: minimum (+) 2. Pain interference: median (+) 3. Number of pain conditions (+)	1. Pain severity: minimum (+) 2. Pain severity: maximum (+) 3. Morning stiffness: median (+)
	D	1. Temperature: mean (+) 2. Dewpoint temperature: mean (-) 3. Relative humidity: mean (+)		1. Dewpoint temperature: SD (+) 2. Relative humidity: SD (-) 3. Mood: median (-)	1. Temperature: mean (-) 2. Dewpoint temperature: mean (+) 3. Pain severity: minimum (+)
For each model from an origin cluster to each outcome in a destination cluster, the three coefficients with the largest absolute values (representing the largest predictors of movement) are reported. All predictors are from the origin week. The direction of association is shown: (+) when an increase in the					

predictor is associated with increased probability of the cluster movement, (-) when a decrease in the predictor is associated with increased probability of the cluster movement.

Discussion

The aim of this study was to develop a model to forecast one week ahead pain severity and identify the predictors of movement between clusters. We used multinomial models to predict movement between origin clusters and destination clusters. The first objective was to optimize the tuning parameters of the elastic net penalty used in these models and the second objective was to select the optimal model from those tested by minimising the misclassification rate. Taken together, these objectives identified that the optimal model for Cluster A (no/low pain) used four groups of covariates and did not deselect features from these groups. The optimal model for Cluster B (mild pain) used all five groups of covariates but selected features via the LASSO penalty term. For clusters C and D (moderate and severe pain) the optimal models used four groups of covariates and selected features via the LASSO penalty term. Multinomial models were compared to a benchmark, cluster-only model. Among cluster-only models, the model with Cluster D as an origin cluster had the lowest misclassification error. The multinomial models with the greatest improvement in model performance over the cluster-only model were Cluster A and Cluster D. The third objective was to identify predictors of movement between clusters. Although there was some heterogeneity in predictors, common predictors were identified: measures of pain severity, pain interference, morning stiffness, fatigue, dewpoint temperature, and the number of pain conditions. Weather covariates were predictors of movement from Cluster D but not from other origin clusters. These predictors impact the probability of cluster movement

for trajectories whose predictor values differ from the average (and therefore are not predictors for each individual trajectory within a cluster).

People with chronic pain have reported that a pain forecast would support daily living[14]. To achieve this goal requires predictors of future pain severity to be identified. Previous studies have reported associations between covariates and concurrent pain severity. For example, Hutchings et al.[16] identified factors (e.g., sleep quality) that were concurrently associated with week-to-week changes in pain scores. We have previously identified factors (e.g., pain interference on daily life) that were concurrently associated with day-to-day changes in pain severity (paper in preparation, CLL, Yimer BB., TH, WGD, DMS, JMcB). Schneider et al.[17] described movement between two latent pain states (lower pain and higher pain) at intervals of two hours, and reported concurrent covariates associated with these movements (e.g., emotional distress, physical functioning). These concurrent associations have identified relationships between exposure and outcome data. However, forecasting pain requires that we predict using prior information about the covariates.

Some studies have forecast symptoms in pain-related long-term health conditions. For example, Kalweit et al.[18] forecast disease activity at the next clinical visit among people with rheumatoid arthritis using a neural network. When modelling disease activity on a 28-point scale, they report good results (8% deviation between predicted and observed values). However, the model underestimated the severity of disease flares and overlooked small fluctuations in disease activity. Therefore, their model overestimated the stability of the disease activity. Similarly, the optimal models in our study had greater accuracy in predicting stability between weeks compared to movement between clusters.

In another example, Vodencarevic et al.[19] used predictive models (including logistic regression) to predict the individual probability of a disease flare among patients with rheumatoid arthritis. They used data from 41 patients and achieved good predictive power. However, their predictors included inflammatory markers which require blood tests for data collection. In contrast, our study used remotely collected patient-generated data and could therefore not include these predictors. Further, participants in our study had a range of chronic pain conditions, some of which were not inflammatory pain conditions, and therefore inflammatory markers would not be suitable candidate predictors.

More broadly, prognostic models have been used in many spheres to predict future health states or outcomes[20] and methodological guidance is available for developing such models[20–23]. In prognostic model research, predictors that are associated with outcomes are called prognostic factors[21] and a list of candidate factors can be explored. In our study, we used a list of candidate predictors to forecast movement between clusters, all chosen due to known clinical relationships between the candidate and pain severity. A different methodological approach would be to prune predictors before model construction, for example by excluding candidates that are not significant in univariate analyses[24]. Having identified predictors of cluster movement in the present study, it is recommended to establish these predictors in other datasets through external validation[21] and this would be an area for future work.

The modest improvement in predictive power over a benchmark model suggests that movement between clusters has a large element of random chance. However, a number of considerations may improve the predictive power of the forecast presented in this study.

1. Predicting other outcomes of pain using the data from this study may provide greater improvements in performance. One outcome that may be considered by future work would be fluctuations in pain severity, as people with chronic pain have expressed this as an important outcome of a pain forecast[14]. We have previously (paper in preparation, CLL, Yimer BB., TH, WGD, DMS, JMcB) quantified measures of pain variability, and reported factors (pain interference, fatigue, mood, morning stiffness, wellbeing) associated with day-to-day changes in pain. It would be feasible to use these measures and associations to forecast pain variability as an alternative outcome.
2. Using different measures of daily-collected data may also improve model performance. For example, many candidate predictors in this study used self-reported data. Such data can lead to misclassification bias, especially between patients who interpret the anchor values of variables differently (e.g., 'moderate fatigue' may have different meanings to different participants)[25]. Other measures of these predictors could lower the misclassification error. For example, sleep quality and exercise can be measured using wearable technologies such as fitness trackers and accelerometer devices[26].
3. Examining predictors on an individual level may improve predictive performance. Previous work has identified that there may be heterogeneous associations between predictors (here, the weather) and outcomes (here, concurrent pain)[27]. Although we examined candidate predictors within clusters of pain severity, these predictors may be stronger (or weaker) on an individual level. However, to create an individual-level forecast would require substantial data input from the participant. For example, Voukantsis et al.[11] required users

to provide 100 days of data before predictions of hayfever symptoms could be fully personalised.

4. Non-linear or interaction relationships between candidate predictors and the outcome variable may exist. For example, unpublished work suggests that time spent outside is a mediating factor between weather variables and pain events (CLL, DMS, Yimer BB., Beukenhorst AL.). These associations were not explored in the present study due to the already large number of candidate predictors. However, such associations may be required for improved forecasts. For example, Chmiel et al.[13] forecast events among people with chronic obstructive pulmonary disease and found that a random forest produced better predictive performance than a logistic regression, possibly due to the inclusion of nonlinear and interaction terms.
5. Other forecasting techniques may provide improved results. In this study, multinomial regressions were chosen for their ease of interpretation (i.e., that the probability of each outcome could be provided to end-users). However, the study by Chmiel et al.[13] found that logistic regression performed only mildly better than baseline, whereas random forests performed significantly better. Thus, other forecasting models may provide improvements to the forecasting capability of pain events.

There were a number of strengths to this study. First, the study benefitted from the use of data from a large population-based study in the United Kingdom. Data collected by over 2000 participants over one year allowed investigation of potential predictors. Second, summarising these data in weekly trajectories permitted inclusion into the study following just two weeks of data collection. Third, the use of multinomial regression model with an elastic net penalty attempted to ensure that the model remained

interpretable through ranking predictors for each outcome and explicitly selecting relevant features for retention in the model.

Some limitations of our approach have already been explored above. The limited use of linear terms in multinomial regression may have restricted the predictive power of the forecasting model. Further, only trajectories with complete pain data were used in the forecasting models. Data may have been missing due to factors influenced by pain severity (e.g., pain is too severe to use a smartphone), and the presence of missing data may act as a predictive factor. Finally, no formal sample size calculation was conducted. Although the elastic net penalty and cross-validation were used to reduce the possibility of overfitting, there remained a high proportion of predictors to events, especially when the number of outcome events was small (e.g., movement from Cluster A to Cluster C/D). It is hence possible that the model has been overfit to the data due to the relatively large number of candidate predictors[28]. To test the impact of overfitting, the model could be validated in an external dataset.

Overall, our study forecast movement between pre-defined clusters of pain severity. Our models successively improved the benchmark model, especially in clusters of no/low pain and severe pain. Different candidate predictors were selected for models from each origin cluster, identifying a need to forecast at a subgroup level. Future work should further explore pain forecasting with other outcomes (e.g., pain variability), different data types (e.g., data from wearable technologies), and other modelling techniques (including exploring interaction terms).

Methods

The reporting of this study followed the TRIPOD guidelines[29]. The underlying code for this study is not publicly available but may be made available to qualified researchers on reasonable request from the corresponding author.

Data

Data Source

Study data were from an mHealth study, Cloudy with a Chance of Pain, and have been previously reported in detail [15,30]. Participants were recruited between January 2016 and January 2017 through a smartphone app developed by uMotif, following advertisements on television, social media, and radio. Data collection was between January 2016 and April 2017. Inclusion criteria were that participants were ≥17 years old, lived in the UK, had at least one chronic pain condition, and owned a smartphone. After downloading the smartphone app, participants were asked to provide baseline data: sex (male/female), year of birth, pain condition(s), and site(s) of pain. Participants were then encouraged to provide data through the app, at least once per day for six months, or longer if they wished. There were 10,584 participants who downloaded the app, provided demographic information, and entered at least one report of pain severity. Baseline data for all participants who downloaded the app can be found in supplementary material of Dixon et al.[15].

In this study, the number of site(s) of pain were summarised as 1 (if one site was selected), 2 (if two separate sites were selected), or 3 (if 3 or more separate sites of pain, or the option “Pain at multiple sites”, or the option “Pain all over body” were selected). The number of chronic-pain conditions was a count of the number of selected conditions.

Daily data provided actively by participants consisted of ten items, answered on a 5-point ordinal scale. The ten items, and their anchors, referred to pain severity (1 = no pain, 5 = very severe pain), fatigue (1 = no fatigue, 5 = very severe fatigue), mood (1 = depressed, 5 = very happy), morning stiffness (1 = no stiffness, 5 = very severely stiff), sleep quality (1 = very poor, 5 = very good), waking up feeling tired (1 = not at all tired, 5 = extremely tired), physical activity (1 = no exercise, 5 = 30+ minutes of strenuous exercise), time spent outside (1 = none of the day, 5 = all of the day), pain impact (1 = not at all, 5 = very much), and wellbeing (1 = very unwell, 5 = very well).

Passive data were also collected as follows. Location of the smartphone was collected hourly through the Global Positioning System (GPS). These locations were used to identify the closest weather station at each hour and could adjust for participants who moved between different locations. Weather data (temperature, pressure, relative humidity, dewpoint temperature, and windspeed) at the weather stations were collected and daily weather conditions were calculated as the mean of these hourly reports[15,30].

Data Preparation

Weekly trajectories of pain were available when a participant reported pain severity each day across a Monday–Sunday week. In a previous study (paper in preparation, CLL, DMS, TH, WGD, JMcB), we reported that these weekly trajectories could be described by four clusters. Each cluster was defined by its medoid (i.e., its most central trajectory). In all four clusters, the medoid reported the same pain severity across the week. Cluster A was defined as no/low pain severity, Cluster B as mild pain severity, Cluster C as moderate pain severity, and Cluster D as severe pain severity.

Of the trajectories assigned to clusters, further inclusion criteria were required for this study. First, a trajectory was required to have been preceded by another trajectory that

had also been assigned to a cluster. The clusters that these trajectories had been assigned were the *Origin cluster* and the *Destination cluster.* If multiple consecutive weeks were contributed by a participant, trajectories may be in an *Origin cluster* for one pair and a *Destination cluster* for another pair. Second, for both weeks in a pair, each covariate was required to have at least two days of available data. This requirement ensured that all summary statistics could be calculated. The day of the week that these data were recorded did not need to be the same for each predictor.

Across all included trajectories, the demographic data of the corresponding participants were summarised by calculating the proportion of females, the mean average age, the mean average number of pain conditions, and the mean average number of sites of pain. Trajectories originally assigned to destinations Cluster A (n = 773), Cluster B (n = 3978), Cluster C (n = 4087), and Cluster D (n = 1612), did not meet the inclusion criteria for this study. We used a *t*-test to compare the proportion of trajectories contributed by males and females, and Kolmogorov-Smirnov tests to compare the distribution of the ages, the number of pain conditions, and the number of pain sites corresponding to trajectories included and excluded in this study. Statistically significant differences were reported when $p < 0.05$. Included trajectories were more likely to be contributed by younger participants and those with more pain conditions. Comparing included and excluded trajectories, there was no statistically significant difference between the proportion contributed by females and the number of pain sites by the corresponding participants.

Model Development

Model development was conducted separately for each of the four clusters (i.e., based on the allocation of the trajectory to one of the four clusters in the previous week). The outcome variable was the cluster in which the following week was allocated, with two

exceptions. Only two trajectories transitioned from Cluster A to Cluster D; therefore, the outcomes of Clusters C and D were combined when developing this model. Similarly, only one trajectory transitioned from Cluster D to Cluster A; therefore, the outcomes of Clusters A and B were combined when developing this model.

Candidate Predictors

A maximum of 55 candidate predictors were included in each model, as described below. For longitudinally-collected covariates, predictors were measures of central tendency (mean for continuous data, median for ordinal data), and measures of range (standard deviation (SD) for continuous data, probability of acute change (PAC) and mean square of successive difference (MSSD) for ordinal data). PAC is a measure of the frequency of fluctuations of a symptom and was calculated as the proportion of non-zero changes within one week. MSSD is a measure that captures the severity and frequency of fluctuations within data; it was calculated by squaring absolute changes in data and finding the average means of these squared changes. Both PAC and MSSD have been highlighted as suitable methods for exploring variability in pain severity[31] and have been used to calculate the variability of pain severity in the Cloudy with a Chance of Pain dataset (paper in preparation, CLL, Yimer BB., TH, WGD, DMS, JMcB). All candidate predictors were standardised so that the mean was 0 and the standard deviation was 1.

Five sets of candidate predictors were identified and grouped by ease of data collection using digital health tools (Figure 2). They were:

(1) Sociodemographic data: Sex, age, number of conditions, number of pain sites.

(2) Pain characteristics using data already required for origin cluster membership: maximum pain severity, minimum pain severity, PAC, MSSD.

(3) Passively collected data from the trajectory in the origin cluster, calculated from available data values of wind speed, temperature, dewpoint temperature, pressure, and relative humidity: mean and SD.

(4) Patient-reported data from the trajectory in the origin cluster, calculated from available data values of self-reported variables: median, PAC and MSSD.

(5) Passively collected data from the trajectory in the destination cluster, calculated from available data values of wind speed, temperature, dewpoint temperature, pressure, and relative humidity during the present week: mean and SD. In a useable forecast, these values would be input from the weather forecast. In this model, the values act as a perfect weather forecast.

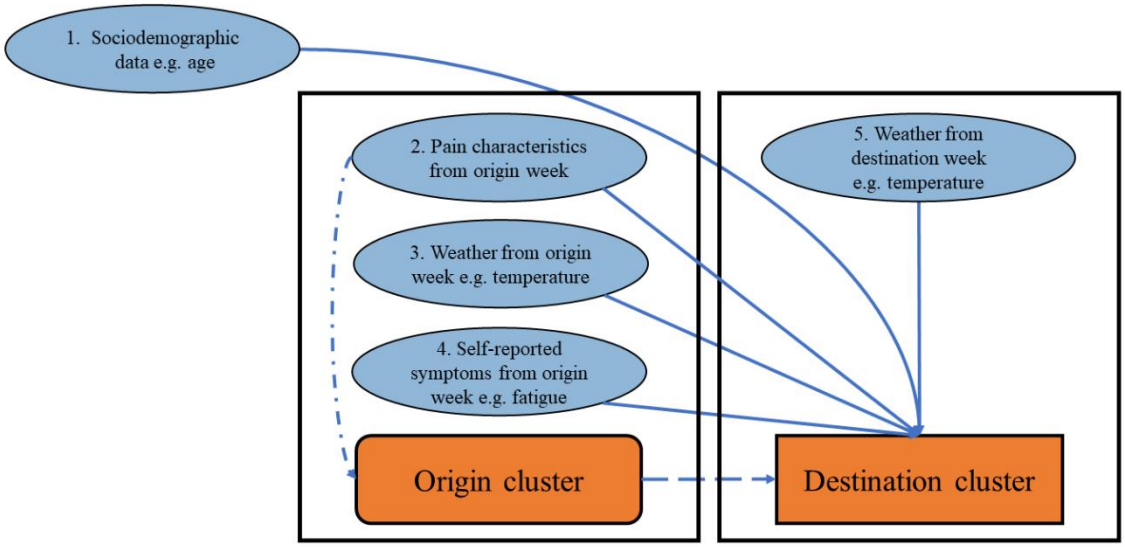


Figure 2: Data used in model development. The origin cluster defines which of the four optimal models are used. A trajectory was assigned to an origin cluster by the self-reported pain severity from previous week. Predictors for the destination cluster are (1) sociodemographic data, (2) pain characteristics of self-reported pain severity from the origin week, (3) weather data from origin week, (4) other self-reported variables from the origin week, and (5) weather data from the destination week.

Candidate models

Candidate models were the same for each cluster with data matrix X. A benchmark model was used to compare the added value of including predictors in the model. The benchmark model chosen was the cluster-only model, which assumed that the next cluster would be the same as the previous cluster, (i.e., 'last value carried forward'[32]). Five candidate models were fit to the data. Each model was a multinomial regression model, with different predictor sets. Of the groups of candidate predictors listed above and in Figure 2, Model 1 used group (1), Model 2 used groups (1) and (2), Model 3 used groups (1–3), Model 4 used groups (1–4), and Model 5 used all 5 groups. The multinomial models were fit using the *glmnet* package[33] in *R* version 4.0.5. Each model

assigned coefficients ($\beta_{j,k}$) to each predictor *j*, for each possible outcome *k*. Therefore, the models estimated either three or four values of $\hat{\beta}_j$ for each predictor (depending on number of outcome levels).

Model Parameters

Statistical model selection involves a compromise between goodness of fit to data and complexity. While adding large numbers of predictors will in general increase the level of agreement between the model and data, this has the risk of overfitting, where observations that are due to random chance or measurement error are misattributed to predictors. A common approach to choosing a model with an optimal number of predictors is to introduce a penalty that grows with the total magnitude of effects inferred and to maximise the combination of the likelihood with this penalty. Popular choices for a penalty include ridge regression[34], which reduces the magnitude of effects inferred without making explicit selection choices and is robust to collinearity among candidate predictors, and the least absolute shrinkage and selection operator (LASSO)[35], which sets some effects to zero and hence selects features by giving them non-zero impacts, but which can be unstable. The elastic net[36] combines the strengths of both the ridge and LASSO while mitigating their limitations.

Estimates of $\hat{\beta}_j$ were calculated by minimising the sum of the negative log-likelihood function and an elastic net penalty term. The elastic net term combines a combination of LASSO (proportion α) and ridge regression ($1 - \alpha$). Therefore, $\alpha = 1$ represents a pure LASSO, and $\alpha = 0$ represents pure ridge regression. The total size of the elastic net penalty term is dictated by the tuning parameter λ. Larger values of λ indicate a larger penalty term, and therefore more shrinkage occurs.

The values of α and λ should be tuned for each model. Eleven values of α were tested, equally spaced between 0 and 1. For each model and for each value of α, an initial run was performed to identify a series of λ values to be tested. Then, for each α and λ pair, we used repeated ten-fold cross-validation to tune α and λ[37]. Due to the reliance on the sampling of the folds, this method was repeated 20 times, so that 20 sets of 10 folds were used. Each time cross-validation was performed, the full data were split into ten subsets, the model was fit for each combination using nine of the subsets, and tested the model on the tenth. This process was repeated ten times, leaving out a different subset each time[36]. The whole process (of splitting the data, training on each combination of nine sets, and testing on the tenth) was repeated 20 times, using the same 20 combinations of subsets of data to test each α and λ pair. The pair of parameters with the lowest average misclassification rate (defined below) were selected as being the optimal parameters for that model.

Model Selection

Models for each origin cluster were selected such that they minimised the misclassification rate. The misclassification rate was calculated as the proportion of weeks assigned by the model to a destination cluster different than the observed destination cluster. Of the five candidate models, the final model was chosen as the one that achieved the lowest misclassification rate. Based on this model, data were assigned probabilities of observing each outcome variable by exponentiating ($\hat{\beta}X$) and normalising so that the sum of probabilities across outcome options equalled 1. Data were assigned to the outcome that reported the highest probability. These predicted outcomes were tabulated in a confusion matrix against the observed outcomes. Numbers on the diagonal of these confusion matrices represented correct classifications.

Predictor Coefficients

Due to the normalisation required, values of $\hat{\beta}$ could not be directly interpreted. However, feature selection via LASSO reduced some $\hat{\beta}_{j,k}$ to zero, and the relative size of the coefficients for each movement between origin and destination clusters could be compared (as covariates had been standardized). Positive coefficients represent an association between an increase in the predictor and an increase in the probability of movement between the corresponding clusters. Negative coefficients represent an association between a decrease in the predictor and an increase in the probability of movement between the corresponding clusters. The three covariates with the largest absolute coefficients were selected for comparison. These covariates have the largest impact on individual trajectories that vary from the cluster average (and so would not impact the probability of all trajectories within the cluster).

Data availability: The datasets generated and analysed during the current study are available from the corresponding author on reasonable request. Data management and analyses were performed in R 4.1.2. Code is available on reasonable request.

Contributions: All authors contributed to the conception of the study. CLL designed and conducted the statistical methods. CLL wrote the first draft of the manuscript. All authors edited the manuscript and agreed to submit for publication.

Acknowledgements: This study was supported by the Centre for Epidemiology Versus Arthritis (grant number 21755). Cloudy with a Chance of Pain was funded by Versus Arthritis (grant reference 21225), with additional support from the Centre for Epidemiology (grants 21755 and 20380). The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Competing interests: Author WGD has received consultancy fees from Google, and author DMS has received consultancy fees from Palta, both unrelated to this work. Authors CLL, TH, and JMcB have no competing interests.

References

1. Vos, T. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* **386**, 743–800 (2015).

2. Cedraschi, C., Luthy, C., Allaz, A. F., Herrmann, F. R. & Ludwig, C. Low back pain and health-related quality of life in community-dwelling older adults. *Eur. Spine J. Off. Publ. Eur. Spine Soc. Eur. Spinal Deform. Soc. Eur. Sect. Cerv. Spine Res. Soc.* **25**, 2822–2832 (2016).

3. Leadley, R. M. *et al.* Healthy aging in relation to chronic pain and quality of life in Europe. *Pain Pract. Off. J. World Inst. Pain* **14**, 547–558 (2014).

4. Stubbs, B. *et al.* Depression and pain: primary data and meta-analysis among 237 952 people across 47 low- and middle-income countries. *Psychol. Med.* **47**, 2906–2917 (2017).

5. Fonseca-Rodrigues, D. *et al.* Correlation between pain severity and levels of anxiety and depression in osteoarthritis patients: a systematic review and meta-analysis. *Rheumatol. Oxf. Engl.* **61**, 53–75 (2021).

6. Sadosky, A. B., DiBonaventura, M., Cappelleri, J. C., Ebata, N. & Fujii, K. The association between lower back pain and health status, work productivity, and health care resource use in Japan. *J. Pain Res.* **8**, 119–130 (2015).

7. Dennis, N. L., Larkin, M. & Derbyshire, S. W. G. 'A giant mess'--making sense of complexity in the accounts of people with fibromyalgia. *Br. J. Health Psychol.* **18**, 763–781 (2013).

8. Flurey, C. A., Morris, M., Richards, P., Hughes, R. & Hewlett, S. It's like a juggling act: rheumatoid arthritis patient perspectives on daily life and flare while on current treatment regimes. *Rheumatology* **53**, 696–703 (2014).

9. Solomon, D. H. & Rudin, R. S. Digital health technologies: opportunities and challenges in rheumatology. *Nat. Rev. Rheumatol.* **16**, 525–535 (2020).

10. Gandrup, J., Ali, S. M., McBeth, J., van der Veer, S. N. & Dixon, W. G. Remote symptom monitoring integrated into electronic health records: A systematic review. *J. Am. Med. Inform. Assoc. JAMIA* **27**, 1752–1763 (2020).

11. Voukantsis, D. *et al.* Personalized symptoms forecasting for pollen-induced allergic rhinitis sufferers. *Int. J. Biometeorol.* **59**, 889–897 (2015).

12. Hurault, G., Domínguez-Hüttinger, E., Langan, S. M., Williams, H. C. & Tanaka, R. J. Personalized prediction of daily eczema severity scores using a mechanistic machine learning model. *Clin. Exp. Allergy J. Br. Soc. Allergy Clin. Immunol.* **50**, 1258–1266 (2020).

13. Chmiel, F. P. *et al.* Prediction of Chronic Obstructive Pulmonary Disease Exacerbation Events by Using Patient Self-reported Data in a Digital Health App: Statistical Evaluation and Machine Learning Approach. *JMIR Med. Inform.* **10**, e26499 (2022).

14. Little, C. L. *et al. What do people living with chronic pain want from a pain forecast? A research prioritisation study*.

http://medrxiv.org/lookup/doi/10.1101/2023.04.24.23289032 (2023) doi:10.1101/2023.04.24.23289032.

15. Dixon, W. G. *et al.* How the weather affects the pain of citizen scientists using a smartphone app. *NPJ Digit. Med.* **2**, 105 (2019).

16. Hutchings, A. *et al.* The Longitudinal Examination of Arthritis Pain (LEAP) study: relationships between weekly fluctuations in patient-rated joint pain and other health outcomes. *J. Rheumatol.* **34**, 2291–2300 (2007).

17. Schneider, S., Junghaenel, D. U., Ono, M. & Stone, A. A. Temporal dynamics of pain: an application of regime-switching models to ecological momentary assessments in patients with rheumatic diseases. *Pain* **159**, 1346–1358 (2018).

18. Kalweit, M. *et al.* Personalized prediction of disease activity in patients with rheumatoid arthritis using an adaptive deep neural network. *PloS One* **16**, e0252289 (2021).

19. Vodencarevic, A. *et al.* Advanced machine learning for predicting individual risk of flares in rheumatoid arthritis patients tapering biologic drugs. *Arthritis Res. Ther.* **23**, 67 (2021).

20. Hemmingway, H., *et al.* Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *BMJ* **346**, e5595 (2013).

21. Riley, R. D., *et al.* Prognosis Research Strategy (PROGRESS) 2: Prognostic Factor Research. *Plos Med.* **10,** e1001380 (2013).

22. Steyerberg, E. W., *et al.* Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. *PLoS Med.* **10**, e1001381 (2013).

23. Higorani, A. D., *et al.* Prognosis research strategy (PROGRESS) 4: Stratified medicine research. *BMJ* **346,** e5793 (2013).

24. Royston, P., Moons, K. G. M., Altman, D. G., Vergouwe, Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ*. **338**, b604 (2009).

25. Kimberlin, C. L. & Winterstein, A. G. Validity and reliability of measurement instruments used in research. *Am. J. Health. Syst. Pharm.* **65**, 2276–2284 (2008).

26. Huhn, S. *et al.* The Impact of Wearable Technologies in Health Research: Scoping Review. *JMIR MHealth UHealth* **10**, e34384 (2022).

27. Yimer, B. B. *et al.* Heterogeneity in the association between weather and pain severity among patients with chronic pain: a Bayesian multilevel regression analysis. *PAIN Rep.* **7**, e963 (2022).

28. Pate, A. et al. Minimum sample size for developing a multivariable prediction model using multinomial logistic regression. *Statistical Methods in Medical Research.* **32**, 555–571 (2023).

29. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med.* **13**, 1 (2015).

30. Schultz, D. M. *et al.* Weather Patterns Associated with Pain in Chronic-Pain Sufferers. *Bull. Am. Meteorol. Soc.* **101**, E555–E566 (2020).

31. Mun, C. J. *et al.* Investigating intraindividual pain variability: methods, applications, issues, and directions. *PAIN* **160**, 2415–2429 (2019).

32. Robert L. McLaughlin. Forecasting models: Sophisticated or naive? *J. Forecast.* **2**, 274 (1983).

33. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, (2010).

34. Hoerl, A. E. & Kennard, R. W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **12**, 55–67 (1970).

35. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **58**, 267–288 (1996).

36. Zou, H. & Hastie, T. Regularization and Variable Selection Via the Elastic Net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67**, 301–320 (2005).

37. Stone, M. Cross-Validatory Choice and Assessment of Statistical Predictions. *J. R. Stat. Soc. Ser. B Methodol.* **36**, 111–133 (1974).

7. Discussion

7.1 Summary of findings

The overall aim of this thesis was to forecast pain among people living with chronic pain. In Section 1.6, I identified four objectives for this thesis, related to the research pipeline described in Section 1.4. I now discuss each of these in turn.

7.1.1 Objective 1

The first objective was to identify an outcome for the pain forecast by conducting work with people living with chronic pain. This was important to ensure that the outcome of the forecast was relevant for people who might use it. In Chapter 3, I reported a focus group and survey that were used to prioritise outcomes of a pain forecast. This study concluded that the timing of flares, the severity of flares, and fluctuations in pain severity were important to people living with chronic pain.

The work in Chapter 3 also reported that 75% of survey respondents would use a short-term pain forecast. This is important because previous work in pain prediction has focussed on long-term outcome predictions (e.g., [109]), but Chapter 3 reports that short-term predictions also have value to people with chronic pain. Specifically, they reported that these predictions would primarily support planning activities, thus improving quality of life. This would support participants who have previously reported frustration in avoiding social commitments [14, 110].

This work informed other chapters and will also inform future directions of work. The importance of pain fluctuations informed Chapter 5, which researched factors that influence daily fluctuations. Future work could further use pain fluctuations as the outcome measure of a pain forecast. Pain flares were discussed by participants in Chapter 3 although further work could define the precise definition of pain flares that is an important outcome to people with chronic pain. Flares can be characterized by increases in pain severity and changes in other symptoms, including fatigue [111–113]. In Chapter 6, I focussed on movement between clusters of pain severity as a first step in exploring these pain flares. Results of transitions to more severe pain categories can be combined to provide probabilities of pain worsening (or pain flares).

Results also identified that understanding triggers of pain would be important, and this indicated that future work should be interpretable, and that predictors of increases in pain should be identified. The focus group within this study also reported concerns which should be addressed in future work. First, data should be collected in such a way as to protect the privacy of the users. Further, the communication of forecasted results should be carefully considered to avoid pain catastrophizing, which can result in increased pain severity [114].

7.1.2 Objective 2

The second research objective was to identify common clusters of pain severity using patient-generated health data. This was important to avoid Simpson's Paradox which suggests that population level associations may mask important group-level associations.

In Chapter 4, I reported cluster analysis on data from Cloudy with a Chance of Pain. Weekly trajectories of data were used, to introduce a social aspect to the data, and to permit users to provide as little as one week of data, or to provide multiple weeks during the analysis. Four clusters of pain severity were identified. The medoids of these clusters were constant throughout the week and were described by the severity of the medoids (no/low pain, mild pain, moderate pain, severe pain).

Section 2.2 reported a systematic search of other studies that examined clusters of pain data. The frequency of data collection ranged in these studies between 2 minutes and 28 years, but among studies of chronic pain (rather than acute, postoperative pain), data were often sparse. The results in Chapter 4 extended current knowledge by examining within-week trajectories of pain severity data, and by quantifying week-to-week transitions between clusters.

Clusters in Section 2.2 were often described by the severity of pain, or by pain improvements or pain worsening. When examining weekly trajectories in Chapter 4, it was unsurprising that longer term trends of improvement or worsening were not identified, but these could be explored by movement between clusters in Chapter 6.

The results from this study informed future chapters in two ways. First, spaghetti plots of the trajectories within each cluster revealed that there remained substantial variability within each cluster that was not described by the cluster medoid. This was further

explored in Chapter 5. Second, a transition matrix identified that 66% of consecutive trajectories remained in the same cluster, but that there was substantial movement between clusters. This suggested the feasibility of using other data to predict movement between clusters. These movements were explored in Chapter 6.

7.1.3 Objective 3

The third research objective was to explore variability of individual trajectories within the clusters identified in Objective 2. This was important as the conclusions of Chapter 3 had identified fluctuations in pain severity as a key outcome of a pain forecast, and increased pain variability is associated with reduced quality of life measures [115, 116] and increased chance of depression [101].

The results section (Quantifying variability) of Chapter 5 reported that around one tenth of trajectories were stable and contained no day-to-day variation. However unstable trajectories contained fluctuations and these were described by three measures. Trajectories in Cluster A (no/low pain) were most likely to be stable and also to have fewer fluctuations within the week. Increases with pain severity were associated with increases in pain impact, fatigue, wellbeing, morning stiffness and mood.

Section 2.3 of this thesis identified that many descriptions of pain variability to date have used implicit measures of variability (i.e., within model variability) or have used standard deviation as a measure of pain variability. Standard deviation was not a suitable measure of pain variability for Cloudy with a Chance of Pain, which collected pain severity data on a 5-point ordinal scale. Instead, other measures suggested by Mun et al. [117] provided a rounded view of the variability seen.

Section 2.3 also reported that variability had been associated in some studies with other factors such as fatigue, cognitive function and sleep quality, but in other studies had not been significantly associated with any variables. The results of Chapter 5 (Identifying variables associated with pain variability) found that some exposures were associated with pain variability in only some clusters, while others (e.g., pain interference) were associated with changes in pain severity across all clusters. Exploring pain variability at a cluster level permitted an understanding of these population-level and group-level associations.

The results of Chapter 5 had implications for future work. First, differences were identified in the variability within different clusters. It was therefore feasible that these fluctuations may also predict movement between different clusters. Hence, variability measures (as well as measures of central tendency) were identified as candidate predictors in Chapter 6. Second, variability had been identified as a potential outcome of a pain forecast in Chapter 3. Therefore, future work should explore whether the variability described in Chapter 6 could be predicted.

7.1.4 Objective 4

The fourth objective was to forecast movement between clusters and identify predictors of this movement. The probability of movement to a more severe pain cluster could be summarised as pain worsening or a pain flare, identified by Chapter 3 as an important outcome of a pain forecast. Further, understanding predictors of this movement would provide information about triggers of pain.

In Section 2.3, I reviewed literature forecasting symptoms of long-term health conditions. I noted that simpler models, such as logistic regression, can have equal predictive power to more complicated models [90]. For this reason, and for model parsimony, I used multinomial regression models to predict movement from *origin* clusters to *destination* clusters in Chapter 6.

In the optimal models identified in this study, the destination cluster was correctly identified in 63.5% – 75.3% of cases. Trajectories in origin Cluster D (severe pain) had the lowest misclassification error. The number of predictors was reduced by using an elastic net penalty and by exploring different groups of candidate predictors. Across the origin clusters, different predictors of movement to each destination cluster were identified, but measures of pain severity, pain interference, morning stiffness, fatigue, dewpoint temperature, and the number of pain conditions were key predictors. Origin Cluster D uniquely identified weather variables in the origin week as predictors for cluster membership in the destination week.

7.2 Strengths and Limitations

There are a number of strengths and limitations to the work conducted throughout this thesis, alongside those described within each chapter.

7.2.1 Strengths

The data analysed throughout this thesis were collected from a longitudinal mHealth study. This data collection benefited from being granular (data were collected daily), and from being collected in participants' own environment. As a result, clusters of pain severity and variability in pain severity could be explored on a granular level.

Section 1.2.2 outlined four challenges in analysing patient-generated health data. These were considered throughout the work conducted in this thesis, to increase the robustness of the results. First, data were treated as ordinal at all stages, to avoid erroneous assumptions about the form of the data. In Chapter 3 (clustering), sensitivity analyses were performed to test whether conclusions would change when data were assumed instead to be continuous or binary. In Chapter 4 (variability), metrics that explicitly assume metric properties of the data such as standard deviation were avoided. In Chapter 5 (forecasting), each cluster was an outcome and assumed to be categorical by the multinomial regression.

Second, the alignment of the data were considered. The data in Cloudy with a Chance of Pain were not aligned to an event such as an operation or the start of a treatment. Instead, weekly trajectories were used throughout the analysis, with each trajectory beginning on a Monday. This introduced a social structure to the data, allowing comparison of trajectories by the data within the trajectory. Chapter 3 further tested this assumption with a sensitivity analysis that instead used data starting on each other day of the week. In these sensitivity analyses, no differences were found between the results among these different trajectories.

Third, the interpretability of the results were considered at each stage. In Chapter 4, clustering was conducted on the values of pain severity explicitly, to ensure that results could be interpreted. Other features of the trajectories (e.g. entropy over time) could have been used to describe the trajectories but would have been less interpretable. In Chapter 5, variability was described using some interpretable measures. Autocorrelation is a measure of the correlation between present and lagged reports. Probability of acute change measures the proportion of changes within the week. In Chapter 6, forecasting was conducted between clusters as this would provide an interpretable outcome. Probabilities of movement to each cluster could be reported to each participant as an interpretable outcome.

Fourth, overfitting was considered in the forecasting model. Overfitting occurs when the model is fit too closely to data and performs poorly in test data. In Chapter 5, the multinomial regressions were fit with cross-validation to minimise the impact of overfitting.

7.2.2 Limitations

One key limitation across the work presented in this thesis was that all analyses required complete weeks of data. Analysing within-week data reduced the length of data collection required, while permitting long-term contributors to provide multiple weeks. However, data missing within-weeks would result in exclusion from analysis. There are reasons why these data may not be missing at random. Participants may be in severe pain and feel unable to use their smartphone for data collection, impacting the proportion of weeks that are assigned to the most severe pain category in Chapter 4. Similarly, participants may have had very low pain and chosen not to enter periods of no pain, impacting the proportion of weeks that are assigned to the no/low pain cluster in Chapter 4. Participants may have consistent pain and may be disengaged by repeating the same data multiple times, which would impact the number of stable trajectories reported in Chapter 5. However, it is possible that these reasons for disengagement would also disengage participants of a pain forecast, and so we can assume that users of a pain forecast may be similar to those reported throughout the chapters of this thesis. Other limitations relating to each chapter were reported in the relevant manuscripts.

7.2 Direction of future work

There are two key directions of future work. The first is to improve the forecasting model, and the second is to create a user-facing mobile app to provide individual results of the forecast.

7.3.1 Improvements to the forecasting model

Chapter 6 forecast movement between clusters as an outcome of multinomial models. This chapter identified factors that were associated with movement between clusters. However, there are a number of improvements that may improve the performance of this model.

First, Simpson's Paradox may be obscuring patterns in the data on an individual level. Clustering aimed to identify associations that could not be observed on the population level. This was seen to be effective in Chapters 5 and 6. In Chapter 5, different associations between exposure data and pain variability were observed among different clusters. In Chapter 5, different predictors were identified within each cluster. Indeed, in Clusters A and D (the more extreme clusters), there was a greater improvement in model performance than in the other clusters. However, clusters were still relatively large and may mask associations that occurred on an individual level. Previous analysis of Cloudy with a Chance of Pain identified that individuals had different (and sometimes opposite) associations between weather exposures and pain events [85], and these could be further exploited. However, this may require a large amount of data from participants before forecasts could be produced. For example, when predicting hayfever symptoms, users were required to provide 100 days of data before individualised forecasts could be produced [108].

Second, many of the candidate predictors were self-reported on a scale that may be interpreted differently between participants. Different interpretations were accounted for in Chapter 5 when changes in self-reported data were examined. In this chapter, changes in symptoms were anchored by the participants own previous reports (e.g., a one-unit increase in fatigue). However, the multinomial model reported in Chapter 6 used candidate predictors with measures of central tendency (mean and median), which may be interpreted differently between participants. As a result, signals in the data may be obscured by different intepretations of self-reporting scales between participants. One solution to this would be to produce individualised forecasts, as discussed above. Another solution would be to use data collected in other ways. For example, wearable technology offers the opportunity to collect sleep quality data or step counts for physical activity, and these could be directly compared between participants. Other studies have used wearable data to understand relationships between pain severity and other factors. For example, Fjeld et al. [66] explored the relationship between physical activity and pain severity using accelerometer data. Using these data, they reported an inverse relationship between step count and pain severity.

Third, other outcome measures could be explored. This thesis (particularly Chapter 6) focussed on movement between clusters as a step towards predicting pain flares. However, other outcome measures may also support people with chronic pain in

understanding future pain. Chapter 3 (Table 3) identified that people with chronic pain were interested in a pain forecast that could predict fluctuations in pain severity. Chapter 5 suggested that this was feasible as there were associations between changes in exposure variables and changes in pain severity. Future work could therefore forecast variability in pain severity.

Fourth, the methods used throughout this thesis aimed to increase the interpretability of results. This was driven by feedback from people with chronic pain who reported that triggers of pain would be a useful feature of a pain forecast (Chapter 3, Table 5). This led to the use of more interpretable models: a probit model to explore variability in Chapter 5 and a multinomial model to forecast movement between clusters in Chapter 6. However, it would be possible to compromise between interpretability and predictive accuracy as there is a growing interest in explainable AI to increase the interpretability of machine learning methods [118].

7.3.2 Adapting a forecast for a mobile app

A pain forecast for predicting movement between pain clusters or predicting other outcomes suggested in Section 7.3.1 could be adapted for use in a mobile app. To provide information to people with chronic pain, measures of the covariates required by the predictive model would be required. Cloudy with a Chance of Pain [82] demonstrated the feasibility of collecting the data required for the forecasting model described in Chapter 6. Further, the model requires only one week (the origin week) of data before predictions could be provided. Each week of new data provided by a participant could produce new predictions for the following destination cluster.

The communication of the predictions to the end-users should be considered. For example, results from Chapter 6 could be communicated as the probability of movement to different clusters in the following week. The key drivers of those predictions could also be communicated to the end-user. However, there are different ways to communicate probabilities, including visual, numerical or verbal measures of uncertainty [119]. Visual representations benefit from concisely reporting information that may otherwise be confusing to the user, but it is generally accepted that these should be augmented with numbers or words to avoid doubt about the outcome [120].

Any such app should be co-developed with end-users. Only 15% of articles reporting an app to support musculoskeletal conditions involved patients in the development of the

app [46]. However, co-development is important to ensure suitability of an app and to avoid research waste [72]. People with chronic pain were involved in the development of the pain forecasting model (Chapter 3), but further work with stakeholders should be conducted in development of a mobile app. For example, this work may include working with stakeholders to discuss the best methods for communication within the app (as discussed above), to discuss the useability of the app, and to discuss ways to minimise anxiety caused by predictions of pain flares (as discussed in the drawbacks in Chapter 3).

An app should take into consideration safety and standards. The focus group described in Chapter 3 highlighted that participants are nervous about the unintended consequences of sharing their data. Another series of focus groups with patients and healthcare professionals reported that participants believed patient generated health data were not fully explored due to a lack of focus on standards [121]. Safety and standards should therefore be at the forefront in the development of a forecasting app.

An app that predicts future pain and identifies triggers of that pain could also support changes in modifiable behaviour. A review found that some apps designed to support people with musculoskeletal conditions promoted physical activity or provided advice on self-management of conditions, and supported this through relaxation therapy [46]. A pain forecasting app could combine education and information about future pain severity to provide targeted interventions. *Just-in-time adaptive interventions* are digital interventions that acknowledge the dynamic nature of symptoms and provide support at the times of highest need [122]. For example, a wearable device detected stress levels in chronic pain patients and provided a mindfulness intervention [123]. This intervention resulted in reduced stressed compared to group-based interventions. A forecasting app could predict increased pain severity (e.g., movement into a worsened pain cluster) and target modifiable covariates associated with this movement.

7.4 Final Remarks

Betty was frustrated at the uncertainty of her pain and wanted to know about her future symptoms. This thesis has described a number of ways that Betty can better understand her present pain and a first step in predicting her future pain. First, we can now tell her that in the current week she will experience one of four pain patterns, each defined by the severity of pain. We can explain that the following week will most likely follow the

same pain pattern, but that a third of transitions will move to other pain patterns. Second, we can describe that her pain is likely to fluctuate during the week, but that these fluctuations are likely to be small. We can tell her that pain changes are associated with pain interference, fatigue, morning stiffness, wellbeing, and mood. As some of these are modifiable factors, Betty could seek support in managing her mood, for example. Finally, we can tell Betty that there are predictors for movement between different pain patterns on consecutive weeks. Predictions are most accurate if she is experiencing no/low pain or severe pain. Future work could also support Betty in understanding future fluctuations in pain and could personalise the predictive model.

8. References

1. Versus Arthritis (2021) Chronic pain in England: Unseen, unequal and unfair. Versus Arthritis

2. Treede R-D, Rief W, Barke A, et al (2019) Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). Pain 160:19–27. https://doi.org/10.1097/j.pain.0000000000001384

3. Fayaz A, Croft P, Langford RM, et al (2016) Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. BMJ Open 6:e010364. https://doi.org/10.1136/bmjopen-2015-010364

4. Vos T, Barber RM, Bell B, et al (2015) Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet 386:743–800. https://doi.org/10.1016/S0140-6736(15)60692-4

5. Barham L (2012) Economic Burden of Chronic Pain Across Europe. J Pain Palliat Care Pharmacother 26:70–72. https://doi.org/10.3109/15360288.2011.650364

6. NHS England Musculoskeletal health. https://www.england.nhs.uk/elective-care-transformation/best-practice-solutions/musculoskeletal/. Accessed 24 Aug 2023

7. Department of Health and Social care working with Public Health England and, Department for Work and Pensions (2019) Musculoskeletal Health: A 5 year strategic framework for prevention across the lifecourse

8. DH/Long Term Conditions (2012) Long Term Conditions Compendium of Information: Third Edition

9. Grant M, O-Beirne-Elliman J, Froud R, et al (2019) The work of return to work. Challenges of returning to work when you have chronic pain: a meta-ethnography. BMJ Open 9:e025743. https://doi.org/10.1136/bmjopen-2018-025743

10. Office for National Statistics (2023) Sickness absence in the UK labour market: 2022

11. Wynne-Jones G, Cowen J, Jordan JL, et al (2014) Absence from work and return to work in people with back pain: a systematic review and meta-analysis. Occup Environ Med 71:448–456. https://doi.org/10.1136/oemed-2013-101571

12. Mutubuki EN, Beljon Y, Maas ET, et al (2020) The longitudinal relationships between pain severity and disability versus health-related quality of life and costs among chronic low back pain patients. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 29:275–287. https://doi.org/10.1007/s11136-019-02302-w

13. Breivik H, Collett B, Ventafridda V, et al (2006) Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment. Eur J Pain 10:287–287. https://doi.org/10.1016/j.ejpain.2005.06.009

14. Flurey CA, Morris M, Richards P, et al (2014) It's like a juggling act: rheumatoid arthritis patient perspectives on daily life and flare while on current treatment regimes. Rheumatology 53:696–703. https://doi.org/10.1093/rheumatology/ket416

15. Fonseca-Rodrigues D, Rodrigues A, Martins T, et al (2021) Correlation between pain severity and levels of anxiety and depression in osteoarthritis patients: a systematic review and meta-analysis. Rheumatol Oxf Engl 61:53–75. https://doi.org/10.1093/rheumatology/keab512

16. Stubbs B, Vancampfort D, Veronese N, et al (2017) Depression and pain: primary data and meta-analysis among 237 952 people across 47 low- and middle-income countries. Psychol Med 47:2906–2917. https://doi.org/10.1017/S0033291717001477

17. Hadi MA, McHugh GA, Closs SJ (2019) Impact of Chronic Pain on Patients' Quality of Life: A Comparative Mixed-Methods Study. J Patient Exp 6:133–141. https://doi.org/10.1177/2374373518786013

18. Hong M, Shin H, De Gagne JC (2019) Social networks, health-promoting behaviors, and health-related quality of life in older adults with and without arthritis. PloS One 14:e0220180. https://doi.org/10.1371/journal.pone.0220180

19. Horng Y-S, Hwang Y-H, Wu H-C, et al (2005) Predicting health-related quality of life in patients with low back pain. Spine 30:551–555. https://doi.org/10.1097/01.brs.0000154623.20778.f0

20. Cedraschi C, Luthy C, Allaz AF, et al (2016) Low back pain and health-related quality of life in community-dwelling older adults. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 25:2822–2832. https://doi.org/10.1007/s00586-016-4483-7

21. Leadley RM, Armstrong N, Reid KJ, et al (2014) Healthy aging in relation to chronic pain and quality of life in Europe. Pain Pract Off J World Inst Pain 14:547–558. https://doi.org/10.1111/papr.12125

22. Crowe M, Whitehead L, Seaton P, et al (2017) Qualitative meta-synthesis: the experience of chronic pain across conditions. J Adv Nurs 73:1004–1016. https://doi.org/10.1111/jan.13174

23. Dennis NL, Larkin M, Derbyshire SWG (2013) 'A giant mess'--making sense of complexity in the accounts of people with fibromyalgia. Br J Health Psychol 18:763–781. https://doi.org/10.1111/bjhp.12020

24. Johnson LM, Zautra AJ, Davis MC (2006) The role of illness uncertainty on coping with fibromyalgia symptoms. Health Psychol 25:696–703. https://doi.org/10.1037/0278-6133.25.6.696

25. Solomon DH, Rudin RS (2020) Digital health technologies: opportunities and challenges in rheumatology. Nat Rev Rheumatol 16:525–535. https://doi.org/10.1038/s41584-020-0461-x

26. De Cock D, Myasoedova E, Aletaha D, Studenic P (2022) Big data analyses and individual health profiling in the arena of rheumatic and musculoskeletal diseases (RMDs). Ther Adv Musculoskelet Dis 14:1759720X221105978. https://doi.org/10.1177/1759720X221105978

27. Gandrup J, Ali SM, McBeth J, et al (2020) Remote symptom monitoring integrated into electronic health records: A systematic review. J Am Med Inform Assoc JAMIA 27:1752–1763. https://doi.org/10.1093/jamia/ocaa177

28. Gandrup J, Selby DA, van der Veer SN, et al (2022) Using patient-reported data from a smartphone app to capture and characterize real-time patient-reported flares in rheumatoid arthritis. Rheumatol Adv Pract 6:rkac021. https://doi.org/10.1093/rap/rkac021

29. Cahn A, Akirov A, Raz I (2018) Digital health technology and diabetes management. J Diabetes 10:10–17. https://doi.org/10.1111/1753-0407.12606

30. Papi E, Murtagh GM, McGregor AH (2016) Wearable technologies in osteoarthritis: a qualitative study of clinicians' preferences. BMJ Open 6:e009544. https://doi.org/10.1136/bmjopen-2015-009544

31. Adler-Milstein J, Nong P (2019) Early experiences with patient generated health data: health system and patient perspectives. J Am Med Inform Assoc 26:952–959. https://doi.org/10.1093/jamia/ocz045

32. Tiase VL, Hull W, McFarland MM, et al (2021) Patient-generated health data and electronic health record integration: a scoping review. JAMIA Open 3:619–627. https://doi.org/10.1093/jamiaopen/ooaa052

33. Check Hayden E (2016) Mobile-phone health apps deliver data bounty. Nature 531:422–423. https://doi.org/10.1038/531422a

34. Shiffman S, Stone AA, Hufford MR (2008) Ecological Momentary Assessment. Annu Rev Clin Psychol 4:1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

35. Lötsch J, Ultsch A (2018) Machine learning in pain research. Pain 159:623–630. https://doi.org/10.1097/j.pain.0000000000001118

36. Bzdok D, Altman N, Krzywinski M (2018) Statistics versus machine learning. Nat Methods 15:233–234. https://doi.org/10.1038/nmeth.4642

37. Hügle M, Omoumi P, van Laar JM, et al (2020) Applied machine learning and artificial intelligence in rheumatology. Rheumatol Adv Pract 4:rkaa005. https://doi.org/10.1093/rap/rkaa005

38. Kao C-K, Liebovitz DM (2017) Consumer Mobile Health Apps: Current State, Barriers, and Future Directions. PM&R 9:S106–S115. https://doi.org/10.1016/j.pmrj.2017.02.018

39. Krebs P, Duncan DT (2015) Health App Use Among US Mobile Phone Owners: A National Survey. JMIR MHealth UHealth 3:e101. https://doi.org/10.2196/mhealth.4924

40. Singh K, Drouin K, Newmark LP, et al (2016) Many Mobile Health Apps Target High-Need, High-Cost Populations, But Gaps Remain. Health Aff Proj Hope 35:2310–2318. https://doi.org/10.1377/hlthaff.2016.0578

41. Dorsey ER, "Yvonne Chan Y-F, McConnell MV, et al (2017) The Use of Smartphones for Health Research: Acad Med 92:157–160. https://doi.org/10.1097/ACM.0000000000001205

42. Grundy Q (2022) A Review of the Quality and Impact of Mobile Health Apps. Annu Rev Public Health 43:117–134. https://doi.org/10.1146/annurev-publhealth-052020-103738

43. Wasil AR, Venturo-Conerly KE, Shingleton RM, Weisz JR (2019) A review of popular smartphone apps for depression and anxiety: Assessing the inclusion of evidence-based content. Behav Res Ther 123:103498. https://doi.org/10.1016/j.brat.2019.103498

44. Mangone ER, Lebrun V, Muessig KE (2016) Mobile Phone Apps for the Prevention of Unintended Pregnancy: A Systematic Review and Content Analysis. JMIR MHealth UHealth 4:e6. https://doi.org/10.2196/mhealth.4846

45. Luo D, Wang P, Lu F, et al (2019) Mobile Apps for Individuals With Rheumatoid Arthritis: A Systematic Review. J Clin Rheumatol Pract Rep Rheum Musculoskelet Dis 25:133–141. https://doi.org/10.1097/RHU.0000000000000800

46. Najm A, Gossec L, Weill C, et al (2019) Mobile Health Apps for Self-Management of Rheumatic and Musculoskeletal Diseases: Systematic Literature Review. JMIR MHealth UHealth 7:e14730. https://doi.org/10.2196/14730

47. Najm A, Lempp H, Gossec L, et al (2020) Needs, Experiences, and Views of People With Rheumatic and Musculoskeletal Diseases on Self-Management Mobile Health Apps: Mixed Methods Study. JMIR MHealth UHealth 8:e14351. https://doi.org/10.2196/14351

48. Thurnheer SE, Gravestock I, Pichierri G, et al (2018) Benefits of Mobile Apps in Pain Management: Systematic Review. JMIR MHealth UHealth 6:e11231. https://doi.org/10.2196/11231

49. Rintala A, Rantalainen R, Kaksonen A, et al (2022) mHealth Apps for Low Back Pain Self-management: Scoping Review. JMIR MHealth UHealth 10:e39682. https://doi.org/10.2196/39682

50. Kristjánsdóttir OB, Fors EA, Eide E, et al (2013) A smartphone-based intervention with diaries and therapist-feedback to reduce catastrophizing and increase

functioning in women with chronic widespread pain: randomized controlled trial. J Med Internet Res 15:e5. https://doi.org/10.2196/jmir.2249

51. Meheli S, Sinha C, Kadaba M (2022) Understanding People With Chronic Pain Who Use a Cognitive Behavioral Therapy-Based Artificial Intelligence Mental Health App (Wysa): Mixed Methods Retrospective Observational Study. JMIR Hum Factors 9:e35671. https://doi.org/10.2196/35671

52. Mork PJ, Bach K, SelfBACK Consortium (2018) A Decision Support System to Enhance Self-Management of Low Back Pain: Protocol for the selfBACK Project. JMIR Res Protoc 7:e167. https://doi.org/10.2196/resprot.9379

53. Rabbi M, Aung MS, Gay G, et al (2018) Feasibility and Acceptability of Mobile Phone–Based Auto-Personalized Physical Activity Recommendations for Chronic Pain Self-Management: Pilot Study on Adults. J Med Internet Res 20:e10147. https://doi.org/10.2196/10147

54. Lo WLA, Lei D, Li L, et al (2018) The Perceived Benefits of an Artificial Intelligence–Embedded Mobile App Implementing Evidence-Based Guidelines for the Self-Management of Chronic Neck and Back Pain: Observational Study. JMIR MHealth UHealth 6:e198. https://doi.org/10.2196/mhealth.8127

55. Zhao P, Yoo I, Lancey R, Varghese E (2019) Mobile applications for pain management: an app analysis for clinical usage. BMC Med Inform Decis Mak 19:106. https://doi.org/10.1186/s12911-019-0827-7

56. Johnson S, Marriott L, Zinaman M (2018) Can apps and calendar methods predict ovulation with accuracy? Curr Med Res Opin 34:1587–1594. https://doi.org/10.1080/03007995.2018.1475348

57. Moglia ML, Nguyen HV, Chyjek K, et al (2016) Evaluation of Smartphone Menstrual Cycle Tracking Applications Using an Adapted APPLICATIONS Scoring System: Obstet Gynecol 127:1153–1160. https://doi.org/10.1097/AOG.0000000000001444

58. De Hond AAH, Leeuwenberg AM, Hooft L, et al (2022) Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. Npj Digit Med 5:2. https://doi.org/10.1038/s41746-021-00549-7

59. DamoLab My Pain Diary. http://mypaindiary.com/. Accessed 15 Jul 2021

60. Dialik Dialik. https://dialik.com/. Accessed 15 Jul 2021

61. Nwankwo T (2014) Hypertension Among Adults in the United States. Med Benefits 31:6

62. Gatchel RJ, Peng YB, Peters ML, et al (2007) The biopsychosocial approach to chronic pain: Scientific advances and future directions. Psychol Bull 133:581–624. https://doi.org/10.1037/0033-2909.133.4.581

63. Dunn KM, Hestbaek L, Cassidy JD (2013) Low back pain across the life course. Best Pract Res Clin Rheumatol 27:591–600. https://doi.org/10.1016/j.berh.2013.09.007

64. Finan PH, Goodin BR, Smith MT (2013) The association of sleep and pain: an update and a path forward. J Pain 14:1539–1552. https://doi.org/10.1016/j.jpain.2013.08.007

65. Kelly GA, Blake C, Power CK, et al (2011) The association between chronic low back pain and sleep: a systematic review. Clin J Pain 27:169–181. https://doi.org/10.1097/AJP.0b013e3181f3bdd5

66. Fjeld MK, Årnes AP, Engdahl B, et al (2023) Consistent pattern between physical activity measures and chronic pain levels: the Tromsø Study 2015 to 2016. Pain 164:838–847. https://doi.org/10.1097/j.pain.0000000000002773

67. Geneen LJ, Moore RA, Clarke C, et al (2017) Physical activity and exercise for chronic pain in adults: an overview of Cochrane Reviews. Cochrane Database Syst Rev 1:CD011279. https://doi.org/10.1002/14651858.CD011279.pub2

68. Kolber MR, Ton J, Thomas B, et al (2021) PEER systematic review of randomized controlled trials: Management of chronic low back pain in primary care. Can Fam Physician Med Fam Can 67:e20–e30. https://doi.org/10.46747/cfp.6701e20

69. Finan PH, Garland EL (2015) The role of positive affect in pain and its treatment. Clin J Pain 31:177–187. https://doi.org/10.1097/AJP.0000000000000092

70. IsHak WW, Wen RY, Naghdechi L, et al (2018) Pain and Depression: A Systematic Review. Harv Rev Psychiatry 26:352–363. https://doi.org/10.1097/HRP.0000000000000198

71. Bair MJ, Robinson RL, Katon W, Kroenke K (2003) Depression and pain comorbidity: a literature review. Arch Intern Med 163:2433–2445. https://doi.org/10.1001/archinte.163.20.2433

72. INVOLVE (2012) Briefing notes for researchers: involving the public in NHS, public health and social care research. INVOLVE, Eastleigh

73. Simpson EH (1951) The Interpretation of Interaction in Contingency Tables. J R Stat Soc Ser B Methodol 13:238–241. https://doi.org/10.1111/j.2517-6161.1951.tb00088.x

74. HAIGH J (2020) MATHEMATICS IN EVERYDAY LIFE. SPRINGER NATURE, Place of publication not identified

75. Maixner W, Fillingim RB, Williams DA, et al (2016) Overlapping Chronic Pain Conditions: Implications for Diagnosis and Classification. J Pain 17:T93–T107. https://doi.org/10.1016/j.jpain.2016.06.002

76. Kongsted A, Kent P, Axen I, et al (2016) What have we learned from ten years of trajectory research in low back pain? BMC Musculoskelet Disord 17:220. https://doi.org/10.1186/s12891-016-1071-2

77. Gooberman-Hill R, Woolhead G, MacKichan F, et al (2007) Assessing chronic joint pain: Lessons from a focus group study. Arthritis Rheum 57:666–671. https://doi.org/10.1002/art.22681

78. Walsh DA, McWilliams DF (2014) Mechanisms, impact and management of pain in rheumatoid arthritis. Nat Rev Rheumatol 10:581–592. https://doi.org/10.1038/nrrheum.2014.64

79. Worley MJ, Heinzerling KG, Shoptaw S, Ling W (2017) Volatility and change in chronic pain severity predict outcomes of treatment for prescription opioid addiction: Volatility and change in pain. Addiction 112:1202–1209. https://doi.org/10.1111/add.13782

80. Bonin Pinto C, Barroso J, Schinitzer TJ (2021) Characterization and implications of daily pain variability and response to treatment in osteoarthritis clinical trials. Osteoarthritis Cartilage 29:S277. https://doi.org/10.1016/j.joca.2021.02.364

81. Friedman J, Hastie T, Tibshirani R (2010) Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw 33:. https://doi.org/10.18637/jss.v033.i01

82. Dixon WG, Beukenhorst AL, Yimer BB, et al (2019) How the weather affects the pain of citizen scientists using a smartphone app. NPJ Digit Med 2:105. https://doi.org/10.1038/s41746-019-0180-3

83. Beukenhorst AL, Schultz DM, McBeth J, et al (2017) Using Smartphones for Research Outside Clinical Settings: How Operating Systems, App Developers, and Users Determine Geolocation Data Quality in mHealth Studies. Stud Health Technol Inform 245:10–14

84. Schultz DM, Beukenhorst AL, Yimer BB, et al (2020) Weather Patterns Associated with Pain in Chronic-Pain Sufferers. Bull Am Meteorol Soc 101:E555–E566. https://doi.org/10.1175/BAMS-D-19-0265.1

85. Yimer BB, Schultz DM, Beukenhorst AL, et al (2022) Heterogeneity in the association between weather and pain severity among patients with chronic pain: a Bayesian multilevel regression analysis. PAIN Rep 7:e963. https://doi.org/10.1097/PR9.0000000000000963

86. Das R, Muldoon M, Lunt M, et al (2023) Modelling and classifying joint trajectories of self-reported mood and pain in a large cohort study. PLOS Digit Health 2:e0000204. https://doi.org/10.1371/journal.pdig.0000204

87. Liddell TM, Kruschke JK (2018) Analyzing ordinal data with metric models: What could possibly go wrong? J Exp Soc Psychol 79:328–348. https://doi.org/10.1016/j.jesp.2018.08.009

88. Walls TA, Schafer JL (2006) Models for intensive longitudinal data. Oxford University Press, Oxford ; New York

89. Rasheed K, Qayyum A, Ghaly M, et al (2022) Explainable, trustworthy, and ethical machine learning for healthcare: A survey. Comput Biol Med 149:106043. https://doi.org/10.1016/j.compbiomed.2022.106043

90. Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 1:206–215. https://doi.org/10.1038/s42256-019-0048-x

91. James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning: with applications in R. Springer, New York

92. Genolini C, Falissard B (2011) Kml: A package to cluster longitudinal data. Comput Methods Programs Biomed 104:e112–e121. https://doi.org/10.1016/j.cmpb.2011.05.008

93. Raykov YP, Boukouvalas A, Baig F, Little MA (2016) What to Do When K-Means Clustering Fails: A Simple yet Principled Alternative Algorithm. PLOS ONE 11:e0162259. https://doi.org/10.1371/journal.pone.0162259

94. Park H-S, Jun C-H (2009) A simple and fast algorithm for K-medoids clustering. Expert Syst Appl 36:3336–3341. https://doi.org/10.1016/j.eswa.2008.01.039

95. Spytkowski M, Kwasnicka H (2012) Hierarchical Clustering through Bayesian Inference. In: Nguyen N-T, Hoang K, Jędrzejowicz P (eds) Computational Collective Intelligence. Technologies and Applications. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 515–524

96. Flynt A, Dean N (2016) A Survey of Popular R Packages for Cluster Analysis. J Educ Behav Stat 41:205–225

97. Collins LM, Lanza ST (2010) Latent class and latent transition analysis: with applications in the social, behavioral, and health sciences. Wiley, Hoboken, N.J.

98. Mo Wang, Bodner TE (2007) Growth Mixture Modeling: Identifying and Predicting Unobserved Subpopulations With Longitudinal Data. Organ Res Methods 10:635–656. https://doi.org/10.1177/1094428106289397

99. Nagin DS, Odgers CL (2010) Group-Based Trajectory Modeling in Clinical Research. Annu Rev Clin Psychol 6:109–138. https://doi.org/10.1146/annurev.clinpsy.121208.131413

100. Zempsky WT, Palermo TM, Corsi JM, et al (2013) Daily changes in pain, mood and physical function in children hospitalized for sickle cell disease pain. Pain Res Manag 18:33–38. https://doi.org/10.1155/2013/487060

101. Schneider S, Junghaenel DU, Keefe FJ, et al (2012) Individual differences in the day-to-day variability of pain, fatigue, and well-being in patients with rheumatic disease: associations with psychological variables. Pain 153:813–822. https://doi.org/10.1016/j.pain.2012.01.001

102. Mun CJ, Suk HW, Davis MC, et al (2019) Investigating intraindividual pain variability: methods, applications, issues, and directions. PAIN 160:2415–2429. https://doi.org/10.1097/j.pain.0000000000001626

103. de Rooij M, van der Leeden M, Heymans MW, et al (2016) Prognosis of Pain and Physical Functioning in Patients With Knee Osteoarthritis: A Systematic Review and Meta-Analysis. Arthritis Care Res 68:481–492. https://doi.org/10.1002/acr.22693

104. Mansell G, Corp N, Wynne-Jones G, et al (2021) Self-reported prognostic factors in adults reporting neck or low back pain: An umbrella review. Eur J Pain Lond Engl 25:1627–1643. https://doi.org/10.1002/ejp.1782

105. Rahman QA, Janmohamed T, Pirbaglou M, et al (2018) Defining and Predicting Pain Volatility in Users of the Manage My Pain App: Analysis Using Data Mining and Machine Learning Methods. J Med Internet Res 20:e12001. https://doi.org/10.2196/12001

106. O'Neill A, O'Sullivan K, McCreesh K (2021) Lower levels of physical activity are associated with pain progression in older adults, a longitudinal study. Eur J Pain Lond Engl 25:1462–1471. https://doi.org/10.1002/ejp.1759

107. Silver JD, Spriggs K, Haberle SG, et al (2020) Using crowd-sourced allergic rhinitis symptom data to improve grass pollen forecasts and predict individual symptoms. Sci Total Environ 720:137351. https://doi.org/10.1016/j.scitotenv.2020.137351

108. Voukantsis D, Berger U, Tzima F, et al (2015) Personalized symptoms forecasting for pollen-induced allergic rhinitis sufferers. Int J Biometeorol 59:889–897. https://doi.org/10.1007/s00484-014-0905-6

109. Nieminen LK, Pyysalo LM, Kankaanpää MJ (2021) Prognostic factors for pain chronicity in low back pain: a systematic review. PAIN Rep 6:e919. https://doi.org/10.1097/PR9.0000000000000919

110. Fullen B, Morlion B, Linton SJ, et al (2022) Management of chronic low back pain and the impact on patients' personal and professional lives: Results from an international patient survey. Pain Pract papr.13103. https://doi.org/10.1111/papr.13103

111. Bykerk VP, Lie E, Bartlett SJ, et al (2014) Establishing a Core Domain Set to Measure Rheumatoid Arthritis Flares: Report of the OMERACT 11 RA Flare Workshop. J Rheumatol 41:799–809. https://doi.org/10.3899/jrheum.131252

112. Berthelot J-M, De Bandt M, Morel J, et al (2012) A tool to identify recent or present rheumatoid arthritis flare from both patient and physician perspectives: The 'FLARE' instrument. Ann Rheum Dis 71:1110–1116. https://doi.org/10.1136/ard.2011.150656

113. Vincent A, Whipple MO, Rhudy LM (2016) Fibromyalgia Flares: A Qualitative Analysis. Pain Med Malden Mass 17:463–468. https://doi.org/10.1111/pme.12676

114. Meints SM, Edwards RR (2018) Evaluating psychosocial contributions to chronic pain outcomes. Prog Neuropsychopharmacol Biol Psychiatry 87:168–182. https://doi.org/10.1016/j.pnpbp.2018.01.017

115. Tupper SM, Rosenberg AM, Pahwa P, Stinson JN (2013) Pain Intensity Variability and Its Relationship With Quality of Life in Youths With Juvenile Idiopathic Arthritis. Arthritis Care Res 65:563–570. https://doi.org/10.1002/acr.21850

116. Zakoscielna KM, Parmelee PA (2013) Pain variability and its predictors in older adults: depression, cognition, functional status, health, and pain. J Aging Health 25:1329–1339. https://doi.org/10.1177/0898264313504457

117. Mun CJ, Suk HW, Davis MC, et al (2019) Investigating intraindividual pain variability: methods, applications, issues, and directions. Pain 160:2415–2429. https://doi.org/10.1097/j.pain.0000000000001626

118. Linardatos P, Papastefanopoulos V, Kotsiantis S (2020) Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy 23:18. https://doi.org/10.3390/e23010018

119. van der Bles AM, van der Linden S, Freeman ALJ, et al (2019) Communicating uncertainty about facts, numbers and science. R Soc Open Sci 6:181870. https://doi.org/10.1098/rsos.181870

120. Spiegelhalter D, Pearson M, Short I (2011) Visualizing Uncertainty About the Future. Science 333:1393–1400. https://doi.org/10.1126/science.1191181

121. Lavallee DC, Lee JR, Austin E, et al (2020) mHealth and patient generated health data: stakeholder perspectives on opportunities and barriers for transforming healthcare. mHealth 6:8. https://doi.org/10.21037/mhealth.2019.09.17

122. Wang L, Miller LC (2020) Just-in-the-Moment Adaptive Interventions (JITAI): A Meta-Analytical Review. Health Commun 35:1531–1544. https://doi.org/10.1080/10410236.2019.1652388

123. Garland EL, Gullapalli BT, Prince KC, et al (2023) Zoom-Based Mindfulness-Oriented Recovery Enhancement Plus Just-in-Time Mindfulness Practice Triggered by Wearable Sensors for Opioid Craving and Chronic Pain. Mindfulness 1–17. https://doi.org/10.1007/s12671-023-02137-0

Appendix A

Appendix to Chapter 1: Submission to Weather, Climate and Society

How Being Inside or Outside of Buildings Affects the Causal Relationship Between Weather and Pain Among People Living with Chronic Pain

Claire L. Little,[a] David M. Schultz,[b] Belay B. Yimer,[a] Anna L. Beukenhorst[c]

[a] *Centre for Epidemiology Versus Arthritis, Manchester Academic Health Science Centre, University of Manchester, Manchester, United Kingdom*

[b] *Centre for Atmospheric Science, Department of Earth and Environmental Sciences, University of Manchester, Manchester, United Kingdom*

[c] *Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, Massachusetts*

For submission as an Article to *Weather, Climate, and Society*

26 June 2023

Corresponding author: Prof. David M. Schultz, david.schultz@manchester.ac.uk

ABSTRACT

Although many people believe their pain fluctuates with weather conditions, both weather and pain may be associated with time spent outside. For example, pleasant weather may mean that people spend more time outside doing physical activity and exposed to the weather, leading to more (or less) pain, and poor weather or severe pain may keep people inside, sedentary, and not exposed to the weather. We conducted a smartphone study where participants with chronic pain reported daily pain severity, as well as time spent outside. We address the relationship between four weather variables (temperature, dewpoint temperature, pressure, and wind speed) and pain by proposing a three-step approach to untangle their effects: (i) propose a set of plausible directed acyclic graphs (DAGs) that account for potential roles of time spent outside (e.g., collider, effect modifier, mediator), (ii) analyze the compatibility of the observed data with the assumed model, and (iii) identify the most plausible model by combining evidence from the observed data and domain-specific knowledge. We find that time spent outside acts as a mediator, explaining 13.9% of the temperature–pain relationship. Thus, an increase in temperature is associated with lower odds of a pain event, partially because of a direct effect and partially because an increase in temperature is also associated with more time spent outside, which in turn is associated with lower odds of pain. Time spent outside also mediates the pressure–pain and wind-speed–pain associations, although to a much lower extent. The other hypotheses either lack support from the data or have assumptions that were unlikely to hold.

SIGNIFICANCE STATEMENT

Three-quarters of people living with chronic pain believe that weather influences their pain. However, people staying inside would not be exposed to weather outside, and good weather may mean that people are more active outside, leading to more or less pain. To obtain data to calculate how the amount of time spent outside affects the weather–pain relationship, we conducted a 15-month smartphone study collecting daily pain reports and nearby weather for nearly 5000 participants. We found that time spent outside mediates the relationship between weather and pain, which helps to explain part of this relationship. Our results show the importance of accounting for other factors when investigating the association between the weather and chronic pain.

1. Introduction

The belief that the weather affects the occurrence and intensity of pain in those people who live with chronic pain is quite old—going back to at least Hippocrates and ancient Chinese medical texts. Three-quarters of people who live with chronic pain believe that their pain levels are influenced by the weather (e.g., Hagglund et al. 1994; Timmermans et al. 2014). Despite this strong anecdotal evidence, however, scientific evidence to support their beliefs has not easily been forthcoming. A review of 43 previously published empirical studies that examined the weather–pain relationship found a wide range of results (Beukenhorst et al. 2020). For example, consider atmospheric pressure. Of the 38 studies reviewed by Beukenhorst et al. (2020) that investigated atmospheric pressure, 20 (53%) reported no link between pressure and pain, 11 (29%) reported high or increasing pressure was associated with higher pain levels, and 7 (18%) reported low or decreasing pressure was associated with higher pain levels. Other weather quantities such as temperature and relative humidity showed similar lack of consensus (Beukenhorst et al. 2020). This lack of scientific consensus leads many doctors to be skeptical about such a relationship, dismissing their patients' observations and concerns about the possible link between their pain and the weather, as well as straining the doctor–patient relationship.

One of the principal reasons for this lack of consensus is the small sample size of these previous studies. Some studies used sample sizes as small as 20 individuals or collected data from the patients over just a few days (Beukenhorst et al. 2020, their Fig. 2). Given that other factors such as physical activity, sleep, and mood are more likely to be associated with people's pain and that weather can exhibit large changes on daily and seasonal time scales, large datasets are needed to tease out the subtle relationships between weather and pain. Modern digital epidemiological approaches such as using smartphone applications (i.e., apps) for more regular reporting and monitoring of people's pain may offer an alternative to traditional approaches in the past (e.g., Wilcox et al. 2012; Jardin et al. 2015; Solomon and Rudin 2020; Beukenhorst et al. 2022), although they are not without their challenges (e.g., De Montjoye et al. 2014; Bol et al. 2018; Mathews et al. 2019; Druce et al. 2019; Onnela 2020; Beukenhorst et al. 2022).

To collect the data to better address the weather–pain relationship, the citizen-science project Cloudy with a Chance of Pain (http://www.cloudywithachanceofpain.com; hereafter, the *Cloudy project*) was created (e.g., Dixon et al. 2019). The Cloudy project was a United Kingdom–based smartphone app study to collect daily self-reported data on

participants' pain. Each day, participants were prompted to answer ten questions about their pain and other measures of well-being (e.g., quality of sleep, mood, physical activity, time spent outside). For example, participants were asked "How severe was your pain today?" and reported pain on a 5-point scale: "no pain," "mild pain," "moderate pain," "severe pain," and "very severe pain." In addition, hourly measurements from the global positioning system (GPS) sensor in the phone (Beukenhorst et al. 2017) allowed the participant's location to be linked to the closest weather station in the Met Office observing network. In this way, we were able to develop daily-averaged weather conditions that each participant experienced, accounting for their travels within the United Kingdom. Collecting these data allowed changes in pain to be linked to weather or changes in weather. Thus, the Cloudy project allowed for creation of a large dataset of participants (over 10,000 participants) with daily records of both pain and weather over a 15-month period.

Our research results from the Cloudy project show that, indeed, our large dataset can produce statistically robust relationships on the effects of weather on people living with chronic pain. Specifically, Dixon et al. (2019) used an epidemiological method called case-crossover to show that an increase in relative humidity, increase in wind speed, and decrease in pressure were associated with higher odds of a pain event in participants. Using synoptic compositing, Schultz et al. (2020) showed that the days with a high percentage of participants who experienced a pain event (defined here as a +1 change or greater in their pain levels on that 5-point scale) were associated with below-normal pressure, above-normal relative humidity, higher precipitation rate, and stronger wind. In contrast, days with a low percentage of participants who experienced a pain event were associated with above-normal pressure, below-normal relative humidity, lower precipitation rate, and weaker wind (Schultz et al. 2020). Yimer et al. (2022) found that between-participant variation existed, both in sensitivity to specific weather conditions and in the direction. Although there was no population-level association between temperature and pain, 1 in 10 participants with chronic pain were sensitive to temperature. Similarly, 1 in 25 were sensitive to relative humidity, 1 in 50 to pressure, and 3 in 100 to wind speed. The Cloudy project continues to yield opportunities for research on the influences on pain in participants.

In particular, one opportunity for deeper understanding of the weather–pain relationship is the self-reported data on the variables other than pain severity, which allow

testing whether time people actually spent outside influences the weather–pain relationship. In particular, two principal issues in using and interpreting this data arise. First, being inside a building may be shelter from some weather conditions, but not others. For example, the temperature and humidity of air inside the building often differs from that outside, and being inside is protection from wind and precipitation. However, air pressure inside a building is generally in equilibrium with the air pressure outside. So, if the weather affects people's pain, such an effect may, for temperature and humidity, be contingent on whether people are outside. Second, weather affects the time people spent outside, as well as the amount of physical activity they do (e.g., Feinglass et al. 2011; Albrecht et al. 2020; Klimek et al. 2022). Specifically, in milder temperatures and in the absence of rain or snow, people spend more time outside and are more physically active. Physical activity may make pain worse or may make pain better, depending on the ailment and the individual. Most participants in our study were living with osteoarthritis or other musculoskeletal diseases (e.g., Dixon et al. 2019), and typically people with these conditions are recommended to do regular low-impact physical activity to reduce pain. Thus, all three variables—weather, time spent outside and physical activity—may influence pain.

Untangling the effects caused by these variables would help clarify the relationship between the weather, pain, and being inside/outside. The Cloudy project team collected information on the relative amount of time that people spent outside. Specifically, participants were asked "How much time have you spent outside today?" They reported back on a 5-point scale: "none," "some," "half," "most," and "all" of the day. Nonetheless, because participants recorded both their pain and time spent outside simultaneously, ascertaining whether pain affects a participant being outside or the weather affects a participant being outside, or both, is not easy to address. For this reason, in the analysis of the Cloudy project data by Dixon et al. (2019) and Schultz et al. (2020), time spent outside was not used as a variable in their analyses. Therefore, this present research aims to further understand the relationship between weather and pain by considering the potential causal relationships between weather, pain and time spent inside/outside.

The remainder of this article is organized as follows. In Section 2, we propose models that explain the possible relationships between weather, pain, and time spent outside for four weather variables: air temperature, dewpoint temperature (a measure of moisture content in the air), atmospheric pressure, and wind speed. These models will be used to

construct statistical tests to evaluate the strength of the relationships between time spent outside, weather, and pain, described in Section 2. Section 3 describes the data and methods of the Cloudy project. Section 4 presents a sequential analysis testing each plausible relationship. Section 5 summarizes the results and concludes the article.

2. Models of the possible ways that time spent outside may affect the causal relationship between weather and pain

Because the Cloudy project data recorded the time spent outside and the weather at the location of the participants, analysis of the results would seem to be straightforward. However, as discussed in Section 1, there may be different ways to describe how time spent outside may affect the causal relationship between weather and pain. (Fig. 1). Here, we illustrate four of those causal relationships through the use of directed acyclic graphs (DAGs; e.g., Tennant et al. 2021).

1. Time spent outside is a *collider* of the weather–pain relationship (Fig. 1a). In this case, both weather and pain affect whether people spent time outside. This situation may be the case where the weather affects people's pain, which then determines whether they spent time outside, as well as the other route where the weather determines the amount of time spent outside.
2. Time spent outside is an *effect modifier* of the weather–pain relationship (Fig. 1b). In this case, the weather outside directly determines people's pain, but there is a different effect when time is spent outside. Therefore, the relationship between weather and pain is different among those participants and days when time is spent outside than when no time is spent outside.
3. Time spent outside is a *mediator* of the weather–pain relationship (Fig. 1c). In this case, weather affects whether people spent time outside (e.g., their exposure to the weather, their level of physical activity), which in turn influences their pain. This situation is in addition to the direct relationship between weather and pain.
4. Time spent outside is a *confounder* of the weather–pain relationship (Fig. 1d). This confound would imply that the time spent outside influences both the weather and people's pain, which would imply that the time spent outside by an

individual influences the weather. This implication is unrealistic, so we reject time spent outside as a confounder in the weather–pain relationship.

Thus, of these four possible models of the effects of time spent outside on the weather–pain relationship, the first three are all plausible. Yet, we do not know what model best explains the Cloudy project dataset. Therefore, we tackle the relationship between four weather variables (temperature, dewpoint temperature, pressure, and wind speed) and pain by proposing a three-step approach to untangle their effects. This approach follows that of Evans et al. (2012). First, we propose a set of plausible directed acyclic graphs that account for potential roles of time spent outside (e.g., collider, effect modifier, mediator), as in Fig. 1. Second, we analyze the compatibility of the Cloudy data with the plausible models. Third, we examine the strengths of the resulting relationships and identify the most plausible model by combining evidence from the observed data and domain-specific knowledge.

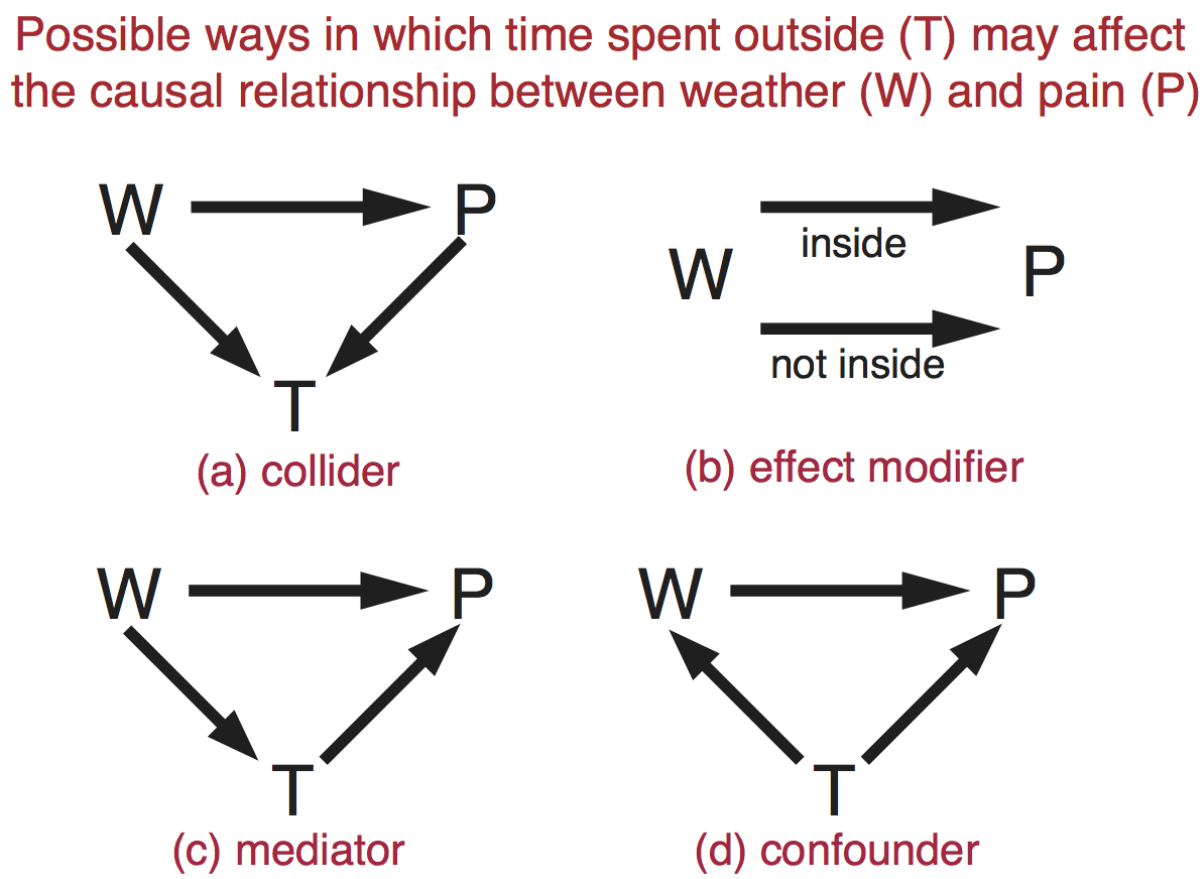


Fig. 1. Directed acyclic graphs of possible ways in which time spent outside (T) may affect the causal relationship between weather (W) and pain (P). (a) T is a collider of the weather–pain relationship, (b) T is an effect modifier of the weather–pain relationship, (c) T is a mediator of the weather–pain relationship, and (d) T is a confounder of the weather–pain relationship. Only (a), (b), and (c) are plausible relationships because time spent outside cannot influence the weather in (d).

3. Data and methods

The description of the dataset largely derives from that of Schultz et al. (2020) as follows in the next five paragraphs.

The study ran for 15 months: 20 January 2016 to 19 April 2017. Before the study commenced, the app was tested and refined through interaction with a patient and public involvement group and a pilot study (Reade et al. 2017). Bolstered by two appearances on British Broadcasting Corporation (BBC) television and other national media attention (Druce et al. 2017; Fig. 2a in Dixon et al. 2019), a total of 13,207 users downloaded the study app during the 12-month recruitment period: 20 January 2016 to 20 January 2017. All 124 United Kingdom postcode areas were represented. Details about the demographics of the population can be found in the supplemental information in Dixon et al. (2019).

A total of 10,584 participants entered their demographic information and at least one pain report, making them eligible for the present study. A total of 6850 (65%) participants remained in the study beyond their first week and 4692 (44%) beyond their first month. Even after 200 days, 15% of participants were still entering data nearly every day (Druce et al. 2017; Fig. 2b in Dixon et al. 2019). This rate of engagement is exceptionally high compared to other mobile health studies where attrition often increases exponentially (e.g., Eysenbach 2005; Druce et al. 2017; Beukenhorst et al. 2022). We believe that our high retention is an indication of the easy-to-use app design, as well as the high level of interest by our participants in contributing toward an answer to this specific research question, which often has been of great personal interest to them (Reade et al. 2017; Druce et al. 2017, 2019).

To examine the relationship between time spent outside, weather, and pain for the Cloudy project dataset, we use the self-reported pain reports from the participants. How to use the pain reports, however, is not straightforward. The self-reported pain levels in our study could lead to bias by directly comparing one participant's pain levels to another. For example, two persons reporting "very severe pain" (5 on our 5-point scale) may have completely different experiences. In addition, participants may use the pain scale differently (e.g., never report maximum or minimum values of pain). To avoid this dilemma, we refer to studies [as reviewed by Olsen et al. (2017)] that show a 20% increase in pain is clinically significant. Such a clinically significant pain event could be measured by an increase of at least 1 category on our 5-point scale. As such, we define a *pain event*

in an individual participant when they report a 1-category or greater increase in their pain level from the previous day (e.g., moderate pain yesterday to severe pain today).

To determine the importance of the relative amount of time spent outside to the weather–pain relationship, we separate the participants who provide a report into those who remain inside (those responding “none” to the question “How much time have you spent outside today?”; INSIDE) and those who spend at least part of the day outside (those responding "some", "half", "most", or "all of the day"; OUTSIDE).

As many as 17% of the participants entered only the demographic information and one day’s report (Druce et al. 2017; Dixon et al. 2019). With no subsequent pain report on the next day, these very low engagers could never experience a pain event in our dataset. We also noticed higher levels of pain reported on days with spikes in recruitment, which was probably an artifact reflecting time needed for participants to settle into their personal scoring system or a “regression to the mean,” where participants joined the study at times of higher than average pain but then settled into a more typical pattern. As such, we excluded data from all participants’ first 10 days in the study.

The dataset used in this article is slightly different than that used in Schultz et al. (2020). Further quality control of the data has been performed, such as incomplete year-of-birth entries being excluded from further analysis. These differences in the datasets are small and do not affect the results.

For each individual, a daily averaged set of weather conditions was calculated. The GPS sensor in the phone was used to determine an individual’s location within the United Kingdom at each hour. In the case of location data being available only once a day, that information is used (Beukenhorst et al. 2017, 2021). That GPS location was linked to the closest weather station in the Met Office observing network. The weather data for the surface weather stations came from the National Oceanic and Atmospheric Administration/National Climatic Data Center Integrated Surface Database (https://www.ncdc.noaa.gov/isd). From these hourly weather reports, an individual daily averaged weather condition was associated with each person on each day. Four weather variables were used in this analysis: 2-m temperature, 2-m dewpoint temperature, sea level pressure (hereafter, pressure), and 10-m wind speed.

Statistics on the resulting dataset are as follows. There were 286,516 participant-days from 4963 participants that were used in this study. Of those, 62,796 were associated with pain events, and 223,720 were not associated with pain events. Of the pain-event

participant-days, 13,852 were associated with the participant being classified as INSIDE, whereas 48,944 were classified as OUTSIDE. Of those participant-days not associated with a pain event, 46,962 were associated with the participant being classified as INSIDE, whereas 176,758 were classified as OUTSIDE. On any given day during the study, a minimum of 15.3% to a maximum of 28.8% of participants had a pain event from among the 130–1267 participants on any given day who entered their pain reports on two consecutive days. The mean rate of pain events for participants who would have entered daily data is 0.311, with a standard deviation of 0.191. Thus, the average person would have 8.7 pain events over a four-week period, if they entered data every day. By choosing "fraction of participants with a +1 increase in pain" as our outcome, our analysis is independent of how an individual interprets the pain scale and implicitly makes a within-person comparison.

4. Results

In this section, we test each of the three plausible models from Section 2 using the statistical methods described below. For each model, we employ a logistic regression. Logistic regressions are often used to fit to data with a binary outcome variable, which in the present study is a pain event. To interpret logistic regressions, odds ratios are reported. For the present study, an odds ratio is the ratio of the probability of a pain event given a certain exposure to the probability of a pain event in the absence of that exposure (e.g., Szumilas 2010). Here, the exposure is the time spent outside. Thus, an odds ratio greater than 1 suggests that the variable is associated with an increased risk of a pain event, whereas an odds ratio less than 1 suggests that the variable is associated with a decreased risk of a pain event. We also report on the confidence intervals at the 95% level. Statistical significance at the 5% level occurs when the confidence interval does not include 1.

a. Is time spent outside a collider of the weather–pain relationship?

If we assume that time spent outside is a collider of the weather–pain relationship (Fig. 1a), then both weather and pain affect an individual's time outside (i.e., $W \rightarrow T$ and $P \rightarrow T$). In this case, it would be erroneous to involve time spent outside in the model, as there is no causal relationship from time spent outside to another variable. In practice, if we assume a collider effect, this is equivalent to examining the relationship between weather and pain as if it were unaffected by time spent outside (i.e., $W \rightarrow P$). In other

words, what would the effect be if we ignored time spent outside and just looked at the relationship between the weather and pain?

To model this relationship under the assumption of a collider, we employ logistic regressions for each weather variable in turn. Each logistic regression is fitted with pain events as the outcome variable and the weather variable as an explanatory variable. Coefficients of the explanatory variables are calculated by the model. We exponentiate these coefficients to provide an odds ratio for each weather variable.

The results of the logistic regression show that temperature, pressure, and wind speed have a small statistically significant relationship with pain events, but that dewpoint temperature does not have a statistically significant relationship with pain events (Table 1). Specifically, an increase in temperature by 1°C reduces the odds of a pain event by 0.3%, an increase in pressure by 1 hPa reduces the odds of a pain event by 0.3%, and an increase in wind speed by 1 m s^{-1} increases the odds of a pain event by 1% (Table 1). For larger changes in the weather, an increase in temperature by 10°C reduces the odds of a pain event by 2.8%, an increase in pressure by 10 hPa reduces the odds of a pain event by 2.5%, and an increase in wind speed by 10 m s^{-1} increases the odds of a pain event by 6.1% (not shown). Although these effects are statistically significant, their small magnitudes mean that these effects may not be clinically meaningful.

Table 1: Effect of the four weather variables on pain by not correcting for time spent outside (i.e., assuming time spent outside is a collider), including the lower and upper bounds on the 95% confidence interval around the odds ratio.

Weather variable	**Odds ratio (95% confidence interval)**
Temperature	0.997 (0.995, 0.999)
Dewpoint temperature	0.999 (0.997, 1.000)

Pressure	0.997 (0.997, 0.998)
Wind speed	1.010 (1.006, 1.015)

If time spent outside is a collider, then collider bias would be introduced by erroneously including time spent outside in the model. This collider bias is the departure of the conditional association of the weather and pain events from their marginal association. We assess this collider bias by examining the difference between the coefficients in the above model (i.e., W $\rightarrow$ P) to those of an alternative model that includes time spent outside as an explanatory variable (i.e., W $\rightarrow$ T). When compared with the odds ratio of the alternative logistic regression that includes time spent outside as an explanatory variable (Table 1), the odds ratios for the weather variables change. Specifically, the odds ratio for temperature declines by 0.0004, the odds ratio for dewpoint temperature declines by 0.0003, the odds ratio for pressure declines by 0.00006, and the odds ratio for wind speed increases by 0.0004, all three being hundredths of a percent changes in odds ratios. Thus, if we incorrectly induced collider bias by adjusting for time spent outside, it would have minimal impact on our findings.

To summarize, three of the four variables (i.e., temperature, pressure, wind speed) show statistically significant relationships with pain events, but these are small effects. If time spent outside were a collider, a model that corrects for that variable would change the relationship, albeit not by much. Thus, we conclude that the collider effect is not a strong effect on an already weak weather–pain relationship.

b. Is time spent outside an effect modifier of the weather–pain relationship?

An effect modifier (Fig. 1b) is equivalent to saying that there is a difference between the relationship between weather and pain events on the days that people spend inside compared to days people spend outside, but that neither the weather nor the pain event would be the cause for people to stay inside in the first place. We can test for the effect modifier by including an interaction term between weather and time spent outside. In practice, because INSIDE/OUTSIDE is a binary variable, the effect modifier is equivalent

to examining two separate W → P relationships, with one being for when the participants are inside (i.e., those responding "none" to the amount of time spent outside) and one for being when the participants are not inside (i.e., those responding "some", "half", "most", or "all of the day"). If these two relationships differ, then the time spent outside is an effect modifier. If these two relationships are the same, then the time spent outside is not an event modifier.

To examine time spent outside as an effect modifier, we employ logistic regression models for each weather variable. In this case, pain event is the outcome variable and there are three explanatory variables: the weather variable, the INSIDE/OUTSIDE variable, and an interaction term between the two. The coefficient of the weather variable is analyzed for those participant-days that were spent entirely inside. The coefficients of all three explanatory variables are summed for those participant-days where some time was spent outside. In both cases, the combination of coefficients is exponentiated to calculate the odds ratio of a pain event.

If we assume that time spent outside is an effect modifier, then the results of this analysis show the following (Table 2). There are no statistically significant differences in the relationships between pressure and wind speed and the odds of a pain event among those days spent inside versus those when some time is spent outside, shown by an overlapping of the confidence intervals in each case. This result is consistent with what we know about pressure being the same inside and outside buildings. Although we do not expect the wind field to affect joint pain directly, changes in wind are often associated with changes in pressure, an interaction that is independent of time spent outside by the participants.

However, there are differences in the relationships between temperature and dewpoint temperature and the odds of a pain event among those days spent inside versus those when some time is spent outside (Table 2). There are positive and significant associations between temperature and dewpoint temperature and the odds of a pain event when no time is spent outside (i.e., INSIDE). In contrast, there are negative and significant associations when some time is spent outside (i.e., OUTSIDE). For temperature and dewpoint temperature, the magnitude of the association between weather and pain is small both for INSIDE and for OUTSIDE, as shown by the lower bound for INSIDE and the upper bound for dewpoint temperature OUTSIDE of the confidence interval being close to 1. This small difference in odds ratios—apparent at the second or third decimal place—is statistically significant, but unlikely to be clinically meaningful.

Thus, time spent outside does not meaningfully modify the relationship between any of the four weather variables and pain. For pressure and wind speed, the difference in the association between those inside and outside was not significant. For temperature and dewpoint temperature, the difference in association between both groups was statistically significant, but small.

Table 2: Effect modifier with 95% confidence intervals in parentheses; the symbol – represents values that are less than 1 at the fourth decimal place, whereas + represents values that are greater than 1 at the fourth decimal place.

Weather variable	Odds ratio for INSIDE (95% confidence interval)	Odds ratio for OUTSIDE (95% confidence interval)
Temperature	1.004 (1.000+, 1.008)	0.996 (0.994, 0.998)
Dewpoint temperature	1.004 (1.000+, 1.008)	0.997 (0.995, 1.000–)
Pressure	0.997 (0.995, 1.000–)	0.998 (0.997, 0.999)
Wind speed	1.011 (1.002, 1.020)	1.010 (1.004, 1.015)

c. Is time spent outside a mediator of the weather–pain relationship?

The mediator effect is explained by comparing the W $\rightarrow$ P relationship to the W $\rightarrow$ T $\rightarrow$ P relationship (Fig. 1c). Hence, testing the mediator effect is only possible for the weather variables that have a statistically significant relationship with pain (Baron and Kenny 1986). In the logistic regressions, pressure, wind speed and temperature are statistically significant with the time spent inside ($p < 10^{-16}$ in each case; not shown). Dewpoint temperature does not have a statistically significant relationship with pain (Table 1) and is therefore excluded from this analysis.

For each model in Table 3, we report the average direct effect (ADE, the direct effect of time spent outside on pain events, W $\rightarrow$ P), the average causal mediation effects (ACME, the indirect effect of the weather variable on the pain event that goes through time spent outside, W $\rightarrow$ T $\rightarrow$ P), and the total effects (ADE + ACME). We then use these quantities to calculate the *percentage mediated* (i.e., the percentage of the weather–pain relationship attributed to time spent outside = ACME $\times$ 100% $\div$ (ADE + ACME)). We

include all these values for completeness, but the most relevant number is the percentage mediated.

The results from these analyses can be interpreted as follows (Table 3). The weather variable remains significant for each of the three outcome models, suggesting that time spent outside only partially mediates the weather–pain relationship. In particular, the proportion of the W $\rightarrow$ P relationship that is mediated by the W $\rightarrow$ T $\rightarrow$ P relationship is 13.4% for temperature (95% CI: 7.2–38.0%), 2.42% for pressure (95% CI: 1.51–4.00%), and 4.14% for wind speed (95% CI: 2.36–8.00%) (Table 3). Thus, participants were most strongly influenced to spend time outside due to the temperature, resulting in a greater likelihood of pain. In contrast, pressure and wind speed were much less likely to influence participants spending time outside, explaining why these have lower percentages mediated. Thus, these results give credence to time spent outside being a mediator for the temperature–pain relationship, and to some smaller extent for the pressure–pain or wind-speed–pain relationships.

Table 3: Testing the mediator effects, with 95% confidence intervals in parentheses, for (a) temperature, (b) pressure, and (c) wind speed. The average direct (ADE), indirect effects (ACME), and the total effect (ADE+ACME) are presented in risk difference scale in line with the mediation R package.

(a) Model	Effect	Odds Ratio (95% confidence interval)
Mediator model	Temperature	0.9615 (0.9596, 0.9632)
Outcome model	Temperature	0.9976 (0.9957, 0.9993)
	Inside	1.0629 (1.0402, 1.0859)
Mediation	ADE	–0.00042 (–0.00073, –0.00011)
	ACME	–0.00007 (–0.00009, –0.00004)
	ADE + ACME	–0.00049 (–0.00082, –0.00016)
	% Mediated	13.9% (7.2%, 38%)

(b) Model	Effect	Odds Ratio (95% confidence interval)
Mediator model	Pressure	0.9942 (0.9934, 0.9950)
Outcome model	Pressure	0.9976 (0.09968, 0.9984)
	Inside	1.0634 (1.0408, 1.0864)
Mediation	ADE	–0.00042 (–0.00055, –0.00029)
	ACME	–0.00001 (–0.00001, –0.00001)
	ADE + ACME	–0.00043 (–0.00057, –0.00030)
	% Mediated	2.42% (1.51%, 4.00%)

(c) Model	Effect	Odds Ratio (95% confidence interval)
Mediator model	Wind speed	1.0413 (1.0368, 1.0458)
Outcome model	Wind speed	1.0099 (1.0055, 1.0143)
	Inside	1.0629 (1.0402, 1.0861)
Mediation	ADE	0.00170 (0.00095, 0.00243)
	ACME	0.00007 (0.00005, 0.00010)
	ADE + ACME	0.00177 (0.00100, 0.00254)
	% Mediated	4.14% (2.36%, 8.00%)

5. Conclusions

Although the belief that the weather affects one's pain has long been held by many people, scientific evidence in support of a consensus has been difficult to achieve, in part because of small data sample sizes required to tease out subtle relationships (e.g., Beukenhorst et al. 2020). These limitations can be overcome through smartphone citizen-science experiments over an extended period of time, as was achieved in Cloudy with a Chance of Pain, which involved over 10,000 participants for a 15-month period (e.g., Dixon et al. 2019). When such limitations are overcome, the popular belief that the weather influences pain events can be confirmed (e.g., Dixon et al. 2019; Schultz et al. 2020; Yimer et al. 2022).

Further analysis of the Cloudy project dataset also shows the value of collecting patient-reported outcomes that help untangle some of the relationships between different weather variables and the time people spend outside. For example, being inside a building is shelter from temperature, dewpoint temperature, and wind, but not pressure. Another example is the causation direction between weather, pain, and time spent outside. If low pain levels caused by favorable weather result in people spending more time outside being physically active and ending up with pain events, then weather is not a simple factor affecting pain only. To understand the possible ways in which time spent outside may affect the causal relationship between weather and pain, we posit four possible relationships for time spent outside: collider, effect modifier, mediator, and confounder. We rule out time spent outside as a confounder since it would imply that the time spent outside by a participant influences the weather. Further, our analysis showed that time spent outside as a collider of the weather–pain relationship is not supported by the data. Evidence to support time spent outside as an effect modifier is weak, with only temperature and dewpoint temperature being significant, but small. Combining domain knowledge and evidence from the observed data, we conclude time spent outside is a mediator of the weather pain relationship and it partially mediates the effect of temperature, pressure, and wind speed on pain. For example, 13.9% of the relationship for temperature–pain is described by people more likely to be inside. This value is considerable, given all the other factors that could affect the weather–pain relationship. Therefore, future researchers should consider this mediating effect.

Untangling the weather–pain relationship is difficult, and our study focused on a subset of possible relationships illustrated through directed acyclic graphs that capture the interplay between weather, time spent outside, and pain. However, there may be other potential directed acyclic graphs involving additional nodes or feedback loops that might further elucidate the relationship. For example, one might consider the application of discovery algorithms to explore other directed acyclic graphs (e.g., Ramsey et al. 2017; Li et al. 2020), but these methods have their own limitations.

If our results are valid and physiological processes can be identified that modulate pain, our results will be important for understanding how patients can learn to deal with chronic pain and its association with weather. These results would suggest approaches for mitigation of pain events in sensitive populations. Specifically, for people whose pain is sensitive to temperature, although the weather has a direct effect on pain, it also influences pain through time spent outside. Thus, by being outside, these individuals may manipulate or counteract part of their pain from the direct effect. In contrast, for people whose pain is sensitive to pressure, there is little difference between being inside or outside.

Acknowledgments. We thank the 10,584 participants in Cloudy with a Chance of Pain who made this study possible with their dedication to daily reporting of their pain levels. We thank Will Dixon for leading Cloudy with a Chance of Pain. We are grateful for the contributions of our patient and public involvement group throughout the study: Carolyn Gamble, Karen Staniland, Shanali Perara, Simon Stones, Rebecca Parris, Annmarie Lewis, Dorothy Slater, and Susan Moore. We thank Bruce Hellman and Ben James at uMotif Limited, London, for their assistance in app design, data collection, and storage. The study app and website were provided by uMotif Limited. The unique flower-like "motif" symptom-tracking interface is owned by uMotif Limited and protected through EU Design Registrations and a US Design Patent. We thank John McBeth, Huai Leng Pisaniello, and Thomas House for their input and comments on earlier drafts of this article. Cloudy with a Chance of Pain was funded by Versus Arthritis (grant reference 21225), with additional support from the Centre for Epidemiology (grants 21755 and 20380). Beukenhorst was supported by a UK Medical Research Council doctoral training partnership (grant MR/N013751/1). Schultz was partially supported by the Natural Environment Research Council UK (grants NE/I005234/1, NE/I026545/1,

NE/N003918/1 and NE/W000997/1). Schultz performed part of this work at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-2210452.

Data Availability Statement. The National Oceanic and Atmospheric Administration/National Climatic Data Center Integrated Surface Database (https://www.ncdc.noaa.gov/isd) provided the weather data from surface stations. Access to the Cloudy with a Chance of Pain dataset is not available at this time.

REFERENCES

Albrecht, B. M., I. Stalling, C. Recke, and K. Bammann, 2020: Accelerometer-assessed outdoor physical activity is associated with meteorological conditions among older adults: Cross-sectional results from the OUTDOOR ACTIVE study. *PLOS ONE,* **15,** e0228053. https://doi.org/10.1371/journal.pone.0228053.

Baron, R. M., and D. A. Kenny, 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology,* **51,** 1173–1182, https://doi.org/10.1037/0022-3514.51.6.1173.

Beukenhorst, A. L., D. M. Schultz, J. McBeth, J. C. Sergeant, and W. G. Dixon, 2020: Are weather conditions associated with chronic musculoskeletal pain? Review of results and methodologies. *PAIN,* **161,** 668–683, https://doi.org/10.1097/j.pain.0000000000001776.

Beukenhorst, A. L., D. M. Schultz, J. McBeth, R. Lakshminarayana, J. C. Sergeant, and W. G. Dixon, 2017: Using smartphones for research outside clinical settings: How operating systems, app developers and user determine geolocation data quality in mHealth studies. *MEDINFO 2017: Precision Healthcare through Informatics,* A. V. Gundlapalli et al., Eds., IOS Press, 10–14, https://doi.org/10.3233/978-1-61499-830-3-10.

Beukenhorst, A. L., J. C. Sergeant, D. M. Schultz, J. McBeth, B. B. Yimer, and W. G. Dixon, 2021: Understanding the predictors of missing location data to inform smartphone study design: Observational study. *J. Med. Internet Res. Mhealth Uhealth*, **9,** e28857, https://doi.org/10.2196/28857.

Beukenhorst, A. L., K. L. Druce, and D. De Cock, 2022: Smartphones for musculoskeletal research – hype or hope? Lessons from a decennium of mHealth studies. *BMC Musculoskelet. Disord.* **23**, 487, https://doi.org/10.1186/s12891-022-05420-8.

Bol, N., N. Helberger and J. C. Weert, 2018: Differences in mobile health app use: A source of new digital inequalities? *The Information Society,* **34,** 183–193, https://doi.org/10.1080/01972243.2018.1438550.

De Montjoye, Y. A., E. Shmueli, S. S. Wang, and A. S. Pentland, 2014: OpenPDS: Protecting the privacy of metadata through SafeAnswers. *PLoS ONE*, **9**, e98790, https://doi.org/10.1371/journal.pone.0098790.

Dixon, W. G., A. L. Beukenhorst, B. B. Yimer, L. Cook, A. Gasparrini, T. El-Hay, B. Hellman, B. James, A. M. Vicedo-Cabrera, M. Maclure, R. Silva, J. Ainsworth, H. L. Pisaniello, T. House, M. Lunt, C. Gamble, C. Sanders, D. M. Schultz, J. C. Sergeant, and J. McBeth, 2019: How the weather affects the pain of citizen scientists using a smartphone app. *npj Digital Medicine*, **2,** 105, https://doi.org/10.1038/s41746-019-0180-3.

Druce, K. L., J. McBeth, S. N. van der Veer, D. A. Selby, B. Vidgen, K. Georgatzis, B. Hellman, R. Lakshminarayana, A. Chowdhury, D. M. Schultz, C. Sanders, J. C. Sergeant, and W. G. Dixon, 2017: Recruitment and ongoing engagement in a UK smartphone study examining the association between weather and pain: Cohort study. *J. Med. Internet Res. mHealth uHealth*, **5**, e168, https://doi.org/10.2196/mhealth.8162.

Druce, K. L., W. G. Dixon, and J. McBeth, 2019: Maximizing engagement in mobile health studies: Lessons learned and future directions. *Rheumatic Disease Clinics of North America,* **45,** 159–172, https://doi.org/10.1016/j.rdc.2019.01.004.

Evans, D., B. Chaix, T. Lobbedez, C. Verger, and A. Flahault, 2012: Combining directed acyclic graphs and the change-in-estimate procedure as a novel approach to adjustment-variable selection in epidemiology. *BMC Med. Res. Methodol.*, **12**, 156, https://doi.org/10.1186/1471-2288-12-156.

Feinglass, J., J. Lee, P. Semanik, J. Song, D. Dunlop, and R. Chang, 2011: The effects of daily weather on accelerometer-measured physical activity. *J. Physical Activity and Health*, **8,** 934–943, https://doi.org/10.1123/jpah.8.7.934.

Hagglund, K. J., W. E. Deuser, S. P. Buckelew, J. Hewett, and D. R. Kay, 1994: Weather, beliefs about weather, and disease severity among patients with fibromyalgia. *Arthritis Rheum.*, **7**, 130–135, https://doi.org/10.1002/art.1790070306.

Jardine, J., J. Fisher, and B. Carrick, 2015: Apple's ResearchKit: Smart data collection for the smartphone era? *J. Roy. Soc. Medicine*, **108,** 294–296, https://doi.org/10.1177/0141076815600673.

Klimek, M., R. S. Peter, M. Denkinger, D. Dallmeier, K. Rapp, D. Rothenbacher, J. Klenk, and the ActiFE Study Group, 2022: The relationship of weather with daily physical activity and the time spent out of home in older adults from Germany–the ActiFE study. *European Rev. Aging and Physical Activity*, **19,** 6, https://doi.org/10.1186/s11556-022-00286-0.

Li, H., Q. Xiao, and J. Tian, 2020: Supervised whole DAG causal discovery. arXiv preprint 2006.04697, https://doi.org/10.48550/arXiv.2006.04697.

Mardini, M.T., S. Nerella, M. Kheirkhahan, S. Ranka, R. B. Fillingim, Y. Hu, D. B. Corbett, E. Cenko, E. Weber, P. Rashidi, and T. M. Manini, 2021: The temporal relationship between ecological pain and life-space mobility in older adults with knee osteoarthritis: A smartwatch-based demonstration study. *JMIR Mhealth Uhealth*, **9,** e19609, https://doi.org/10.2196/19609.

Mathews, S. C., M. J. McShea, C. L. Hanley, A. Ravitz, A. B. Labrique, and A. B. Cohen, 2019: Digital health: A path to validation. *npj Digital Medicine*, **2,** 38, https://doi.org/10.1038/s41746-019-0111-3.

Onnela, J.-P., 2020: Opportunities and challenges in the collection and analysis of digital phenotyping data. *Neuropsychopharmacology*, **46,** 45–54, https://doi.org/10.1038/s41386-020-0771-3.

Ramsey, J., M. Glymour, R. Sanchez-Romero, and C. Glymour, 2017: A million variables and more: The Fast Greedy Equivalence Search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images. *International Journal of Data Science and Analytics*, **3,** 121–129. https://doi.org/10.1007/s41060-016-0032-z.

Reade, S., K. Spencer, J. C. Sergeant, M. Sperrin, D. M. Schultz, J. Ainsworth, R.

Lakshminarayana, B. Hellman, B. James, J. McBeth, C. Sanders, and W. G. Dixon, 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. *J. Med. Internet Res. mHealth uHealth*, **5**(3), e37, https://doi.org/10.2196/mhealth.6496.

Schultz, D. M., A. L. Beukenhorst, B. B. Yimer, L. Cook, H. L. Pisaniello, T. House, C. Gamble, J. C. Sergeant, J. McBeth, and W. G. Dixon, 2020: Weather patterns associated with pain in chronic-pain sufferers. *Bull. Amer. Meteor. Soc.*, **101,** E555–E566, 675–678, https://doi.org/10.1175/BAMS-D-19-0265.1.

Solomon, D. H., and R. S. Rudin, 2020: Digital health technologies: Opportunities and challenges in rheumatology. *Nature Reviews Rheumatolology*, **16,** 525–535, https://doi.org/10.1038/s41584-020-0461-x.

Szumilas, M., 2010: Explaining odds ratios. *J. Can. Acad. Child Adolesc. Psychiatry*, **19,** 227–229; Erratum, **24,** 58.

Tennant, P. W., E. J. Murray, K. F. Arnold, L. Berrie, M. P. Fox, S. C. Gadd, W. J. Harrison, C. Keeble, L. R. Ranker, J. Textor, and G. D. Tomova, 2021: Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: Review and recommendations. *International Journal of Epidemiology*, **50,** 620–632, https://doi.org/10.1093/ije/dyaa213.

Timmermans, E. J., and Coauthors, 2014: 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 Musculoskeletal Disord.*, **15**, 66, https://doi.org/10.1186/1471-2474-15-66.

Wilcox, A. B., K. D. Gallagher, B. Boden-Albala, and S. R. Bakken, 2012: Research data collection methods: From paper to tablet computers. *Medical Care*, **50,** S68–S73, https://www.jstor.org/stable/23216661.

Yimer, B. B., D. M. Schultz, A. L. Beukenhorst, M. Lunt, H. L. Pisaniello, T. House, J. C. Sergeant, J. McBeth, and W. G. Dixon, 2022: Heterogeneity in the association between weather and pain severity among patients with chronic pain: A Bayesian multilevel regression analysis. *PAIN Reports*, **7,** e963, https://doi.org/10.1097/PR9.0000000000000963.

Appendix B

Appendix to Chapter 2: Table of papers from literature search in Section 2.2

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Hundert, 2022, Canada [1]	Infant pain following heel lance using secondary analysis of RCT of pain control	236 participants; Age: 32.9 weeks; 44.1% female	7 measures on 4-point scale	4 times over 2 minutes (30, 60, 90, 120 seconds)	Group-based trajectory modelling	Not mentioned	BIC, posterior probability >0.7, class size	Five: not described
Kannampallil, 2016, USA [2]	General pain in inpatients using primary data of observational study	7762 participants; Age: 43.3; 64.5% female	11-point NRS recorded by nurses	Average 13.45 times over 48 hours (as many times as recorded)	*k*-means clustering from a mixed-effects regression model	All available data used	Scree plot, AUC	Four: Not explicitly described
Weng, 2020, USA [3]	Post-operative pain after TKA using secondary analysis of a database	3442 participants; Age: 67.7; 62.2% female	11-point NRS	Median 22 times over 3 days (as much as available)	*k*-medians using linear regression results	At least 10 data points required	Between cluster variation ≥ 5%, group size ≥ 5%	Four: slightly rise, completely drop, sudden rise, steady

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Page, 2019, Canada [4]	Post-operative pain after hepatic resection using primary data of observational study	152 participants; Age: 63; 41% female	11-point NRS recorded at rest and coughing by nurse	Every 8 hours for 48 hours & daily until discharge	Growth mixed modelling	Not mentioned	AIC, BIC, group size ≥ 5%, theoretical soundness, interpretability	Four: Constant mild pain, constant mild/moderate pain, constant moderate pain, severe pain
Wu, 2022, Taiwan [5]	Acute pain in emergency department using secondary analysis of hospital database	28105 participants; Age: 63; 45% female	11-point NRS	Maximum 72 reports over 72 hours (hourly)	Group-based trajectory modelling	At least three data points required	BIC	Three: no pain, moderate-to-severe pain fast resolvers, moderate pain slow resolvers
Thomazeau, 2016, France [6]	Chronic postsurgical pain after TKA using primary data from an observational study	109 participants; Age: 69.2; 28.4% female	11-point NRS at rest and on movement	Average 3 times over 4 days	Latent class growth analysis	Only complete data used	AIC, BIC, entropy, BLRT, group size > 25	Two: low-intensity, high-intensity
Awadalla, 2022, USA [7]	Postoperative pain using primary data from observational study	2106 participants; Age: 62; 59% female	11-point NRS	At least 4 times in 96 hours	Feature selection, PCA, *k*-means	At least four data points required	Cubic clustering criterion	Four: undescribed
Althaus, 2018, Germany [8]	Postoperative pain using primary data of observational study	174 participants; Age: 51.91; 44% female	11-point NRS of pain on movement	5 times over 5 days (daily)	Growth mixture modelling	Only complete data used	BIC, LMR-LRT	Three: High initial pain with high resolution, high initial pain with moderate resolution, low initial pain with moderate resolution

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Getachew, 2023, Ethiopia [9]	Postsurgical pain using primary data from observational study	218 participants; Age: 33.4; 19.3% female	11-point NRS	5 times over 5 days (daily)	Growth mixture modelling	Not mentioned	Not reported	Two: rapid pain relief, consistently high pain
Ilhan, 2020, Australia [10]	Post-surgical pain in infants using data from retrospective medical record review	726 participants; Age: 36.0 weeks; 46.3% female	Pain Assessment Tool (range 0–20)	Variable (on admission, on return to ward after surgery, 2 hour for 24 hours, 4 hours until 24 hours of no analgesia)	Group-based trajectory analysis	Unclear due to data collection	Clinically interpretability, parsimony	Two: Typical pain trajectory, atypical pain trajectory
Page, 2016, Canada [11]	Hip pain after THA using secondary analysis of a randomized control trial	150 participants; Age: 60.0; 48% female	Presence of chronic pain at surgery site, 6-point NRS, pain frequency	17 times over 5 days (4, 8, 12, 16, 24, 32, 40, 48, 56, 64, 72, 80, 88, 96, 104, 112, 120 hours)	Growth mixture model	Not mentioned	BIC, interpretability, parsimony, group size ≥ 5%	Four: Low-to-moderate followed by a steady decline; low-to-moderate followed by consistent pain; severe pain followed by quick and steady decrease; moderate pain followed by steady pain
Ocay, 2020, Canada [12]	Postoperative pain after spinal fusion surgery using primary data from an observational study	106 participants; Age: 15.4; 76.4%	11-point NRS reported by the nurse	Average 20 times over 120 hours (On average: every 6 hours)	Growth mixed models	All available data used	AIC, BIC, entropy, group size ≥ 5%, parsimony	Four: constant mild, constant mild-to-moderate, mild-to-moderate increasing, constant moderate

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Thomas, 2020, UK & Netherlands [13]	Knee pain using secondary analysis of a randomized control trial	449 participants; Age: 45.0; 41% female	11-point NRS on average daily pain	6 times over 6 days (daily)	Latent curve growth analysis	Only complete data used	AIC, BIC, ABIC, entropy, group size ≥ 5%, posterior probabilities > 0.7	Six: determined by rate of decrease in pain and final pain
Mamoun, 2022, USA [14]	Postoperative pain using primary data	1660 participants; Age: 62.2; 56.0% female	11-point NRS	10 times over 6 days (0, 6, 12, 24, 36, 48, 60, 72, 96, 120, 144 hours)	Latent class linear mixed models	Not mentioned	BIC, class size >100	Four: stationary pain, rapidly improving pain, slowly improving pain, acute worsening pain
Chang, 2022, Taiwan [15]	Postoperative pain using primary data from a retrospective study	635 participants; Age: 63; 56.2% female	11-point NRS	7 times over 7 days (daily)	Group-based trajectory modelling	Models fit to all available data	BIC	Two: Mild pain, rebound pain
Vasilopoulos, 2021, USA [16]	Postoperative pain using primary data from prospective cohort study	360 participants; Age: 59; 50% female	11-point NRS	7 times over 7 days (daily)	Group-based trajectory modelling	Models fit to all available data	BIC, Bayes factor, cluster size > 1%	Five: moderate-high, moderate-low, high, low, decreasing
Jaffa, 2021, USA [17]	Subarachnoid haemorrhage with secondary analysis of observational database	305 participants; Age: 55.8; 66.9% female	11-point NRS	13 times over 13 days (daily)	Group-based trajectory modelling	Not mentioned	BIC, AIC, parsimony	Five: pain free, low pain intermittent spikes slight increase, increasing pain then mild improvement, maximum pain steady decrement, moderate pain subtle improvement

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Daoust, 2019, Canada [1F8]	Acute pain in the emergency department when discharged with opioids using primary data from a prospective cohort study	372 participants; Age (median): 54; 50%	11-point NRS of daily average pain	14 times over 2 weeks (daily)	Group based trajectory modelling	Maximum likelihood estimations	BIC, Bayes factor, group size ≥ 5%, posterior probabilities	Six: Severe-severe, severe-moderate, severe-mild, severe-no pain, moderate-mild, mild-no pain
Cummings, 2022, USA [19]	Pain in adolescents using secondary data from observational study	155 participants; Age: 16; 49.68% female	101-point NRS	10 times over 10 days (daily)	Latent class analysis (of slopes and intercepts from mixed-effects linear regression)	At least five data points required	Entropy (>80%), class size >5%, LMR LRT, interpretability	Three: Slow decreases in pain, rapid decreases in pain, stable or slight increases in pain
Vowles, 2017, UK [20]	Chronic pain (receiving Acceptance and Commitment Therapy) using primary data in observational study	174 participants; Age: 45.9; 69.9% female	11-point NRS	4 times over 4 weeks (weekly)	Growth mixture models	Maximum likelihood estimation	BIC, aBIC, VLMR-LRT, entropy, inspection	One: homogeneous group
Walton, 2014, Canada [21]	Neck pain using primary data of observational study	50 participants; Age: 39.9; 76.5% female	11-point NRS for worst pain	4 times over 4 weeks (weekly)	Latent Class Growth Analysis	All participants included	BIC, BLRT, entropy	Three: worsening, rapid improvement, slow improvement
Mori, 2021, USA [22]	Postoperative pain using primary data of observational study	75 participants; Age: 64; 24% female	11-point NRS	10 times over 30 days (every 3 days)	Group-based trajectory modelling	At least three data points required	BIC, posterior probabilities	Four: persistently low, moderate declining, high declining, persistently high

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Connor, 2020, USA [23]	Pain following traumatic experience using secondary analysis of observational study	87 participants; Age: 35.02; 37.5% female	11-point NRS for daily pain	36 times over 36 days (daily)	Latent class growth analysis	Averaged over 4 days to account for missing data	BIC, adjusted BIC, AIC, entropy, VLMR-LRT, BLRT	Three: low, high, descending
Shearer, 2022, Canada [24]	Cerebral palsy in children with primary data from cohort study	101 participants; Age: 12.9; 48.5% female	11-point scale using faces	5 times over 5 weeks (weekly)	Latent class growth modelling	Few missing data so no imputation	BIC, AIC, parsimony	Five: very low, low, high, moderate changing, high decreasing to moderate
Singh, 2019, USA [25]	Postoperative pain after total knee arthroplasty using secondary analysis of observational study	659 participants; Age: 67.1. 64.5% female	Time point one: KOOS. Subsequent time points: 11-point NRS of pain.	3 times over 8 weeks (preoperative, 2 weeks, 8 weeks)	Group-based trajectory models	Imputed using predicted values from linear regression models	BIC, Adjusted R^2, Hosmer-Lemeshow test, clinical interpretability	Two: Fast pain response, slow pain response
Downie, 2016, Australia [26]	Acute lower back pain using secondary analysis of randomized control trial	1585 participants; Age: 45.15; 46.7% female	11-point NRS	5 times over 12 weeks (Baseline, 1, 2, 4 and 12 weeks)	Latent class growth analysis	Excluded if missing > 2 points. Maximum likelihood estimation.	BLRT, AIC, BIC, entropy, posterior probabilities, group size, distinct trajectories	Five: rapid recovery, recovery by week 12, incomplete recovery, fluctuating pain, persistent high pain

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Lee, 2018, USA [27]	Knee OA using primary data of a randomized control trial	171 participants; Age: 61; 71% female	WOMAC score to measure pain intensity	12 times over 12 weeks (weekly)	Group-based trajectory modelling	Assumed missing at random. Analysis performed on available data.	BIC, group size (≥ 10%)	Four: lower pain early improvement, moderate pain early improvement, higher pain delayed improvement, higher pain no improvement
Daoust, 2013, Canada [28]	Post-operative after minor thoracic injury using primary data of prospective, multicentre study	734 participants; Age: 54; 36.5% female	11-point NRS	5 times over 3 months (Hospital visit and 1 week, 2 week, 30 day, 90 day later)	Latent class growth modelling	Only patients with ≥ 3 time points included	BIC, posterior probabilities, group size ≥ 5%, clinically significant pain (>3/10)	Three: low pain then no clinically significant pain, moderate pain then no clinically significant pain, clinically significant pain at end
Atukorala, 2022, Australia [29]	Knee OA using secondary analysis of observational study	313 participants; Age: 62.2; 60.4% female	KOOS-p: 5-point Likert scale	9 times over 3 months (every 10 days)	Latent variable longitudinal mixture models	Estimated using maximum likelihood	AIC, aBIC, VLMR-LRT, entropy, posterior probability, interpretation	Three: low-moderate pain with large improvement, minimal change, moderate-high pain with worsening
Burns, 2020, USA [30]	Acute pain at emergency department using primary data of prospective study	375 participants; Age: Not recorded; 100% female	11-point NRS	4 times over 4 months (monthly)	Latent class growth analysis	All available data used	BIC, visual inspection, posterior probabilities ≥ 0.7	Three: Early recovery, delayed recovery, no recovery

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Braunwalder, 2022, Switzerland [31]	Spinal cord injury using secondary analysis of cohort study	343 participants; Age: 53.5; 25.9% female	11-point NRS	4 times over at least 20 weeks (4, 12, 24 weeks, discharge)	Latent process mixed modelling	At least two data points required. Sensitivity analysis with imputed values.	AIC, BIC, aBIC, entropy, class size >5%	Four: stable moderate, decreasing, increasing, stable low
Toyoda, 2017, Japan [32]	Back pain after acute osteoporotic vertebral fractures using primary data of prospective multicentre cohort study	128 participants; Age: 78.0; 83.6% female	VAS 0–100mm based on average back pain in the past week	4 times in 6 months (baseline and 1, 2, 6 months)	Hierarchical cluster analysis	Not mentioned	Subjective choice using dendrogram	Four: Average, excellent, fluctuating, persistent severe
Schuller, 2021, Netherlands [33]	Low back pain using primary data from prospective cohort study	1117 participants; Age: 47.5; 60.0% female	11-point NRS	5 times over 6 months (every six weeks)	Latent class growth analysis	Sensitivity analyses, multiple imputation	VLMR LRT, BIC, interpretability, clinical practicality, class size	Three: Improved, Not improved, Low baseline
Dunn, 2006, UK [34]	Lower back pain using primary data of observational study	342 participants; Age: 30-59; 58% female	11-point NRS for least and usual pain over two weeks and current pain	6 times over 6 months (monthly)	Longitudinal latent class analysis	Individuals removed if missing more than three data points	AIC, BIC, CAIC, model fit likelihood ratio chi squared statistic, bootstrap *p*-value	Four: Persistent mild, recovering, severe-chronic, fluctuating

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Axen, 2011, Sweden [35]	Lower back pain using primary data of observational study	165 participants; Age: 45; 46% female	8-point NRS	26 times over 6 months (weekly)	Hierarchical clustering	Included only those with ≥ 80% data	Calinski-Harabasz	Four: Stable, fast improvers, slow improvers, typical patient
Ailliet, 2018, Belgium & Netherlands [36]	Lower back pain & neck pain (with chiropractic treatment) using primary data of prospective cohort study	448 participants; Age: 40.4; 48.7% female	11-point NRS	26 times over 26 weeks (weekly)	Latent class growth analysis	Not mentioned	BIC, posterior probabilities, parsimony	Four for neck pain: recovering from mild, recovering from high, severe-chronic, recovering from mild with a flare up. Four for back pain: recovering from mild, recovering from high, moderate-chronic, slowly recovering from high
Tamcan, 2010, Switzerland [37]	Lower back pain using primary data of cross sectional study	305 participants; Age: 53.1; 53% female	7-point NRS	26 times for 26 weeks (weekly)	Latent class analysis	Included participants with 27 data points, gaps no longer than 3 weeks	BIC, AIC, L2 reduction from cluster one, comparison to previous study	Four: severe persistent, moderate persistent, mild persistent, fluctuating

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Brennan, 2020, USA [38]	Pain in nursing home patients using analysis of nursing home records	2539 participants; Age: 73.0; 3.6% female	11-point NRS or 4-point verbal pain	4 times over 9 months (baseline, 3, 6, 9 months)	Growth mixed models	Only complete data used	Successful model convergence, LRTs, AIC, BIC, entropy, elbow tests, class size, substantive meaning	Four: high-increasing, moderate-increasing, low-increasing, low-decreasing
Dumenci, 2019, USA [39]	Knee pain following TKA using secondary analysis of a randomized trial	384 participants; Age: 63.2; 67% female	WOMAC score for pain on exercise	4 times over 12 months (2 weeks pre-surgery and 2, 6, 12 months post-surgery)	Piecewise latent class growth analysis	Full information maximum likelihood method	AIC, BIC, adjusted BIC, RMSEA, Tucker-Lewis index, LRT	Two: Good recovery and poor recovery
Ellyson, 2022, USA [40]	Chronic postsurgical pain in adolescents with primary data from cohort study	117 participants; Age: 14.5; 64.1% female	11-point NRS averaged over 7 days	4 times over 12 months (Baseline, 2 weeks, 4, 12 months	Group-based trajectory modelling	Full information maximum likelihood	AIC, BIC, entropy	Two: Declining, high and persistent
Lee, 2020, USA/ Canada [41]	Musculoskeletal trauma using secondary analysis of two cohorts	205 participants; Age: 39.7; 45.1% female	11-point NRS	4 times over 12 months (baseline, 2-4 weeks, 3 months, 12 months)	Latent class growth analysis and *k*-means	Only included if ≥ 2 data points	AIC, BIC, entropy, group size, parsimony, LMR-LRT	Four: rapid recovery, minimal or no recovery

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Muller, 2021, Switzerland [42]	Postoperative low back pain using primary data	141 participants; Age: 61.1; 57% female	11-point NRS	4 times over 12 months (Baseline, 2, 6, 12 months)	Group-based trajectory modelling	Multiple imputation and sensitivity analyses	AIC	Three: complete, incomplete, or no recovery
Panken, 2016, Netherlands [43]	Lower back pain using secondary analysis of three randomized control trials	240 participants; Age: 40.6; 28.9% female	VAS	4 times over 1 year (3 monthly)	Latent class growth analysis	All available information used	BIC, BLRT, entropy, posterior probabilities, clinical relevance	Three: chronic high persistent pain, recovering chronic pain, mild persistent pain
Rabbitts, 2015, USA [44]	Post-operative pain in children using primary data of observational study	60 participants; Age: 14.7; 66.7% female	11-point NRS for intensity, 5-point NRS for frequency and 4-point NRS for duration	4 times over 1 year (baseline, 2 weeks, 4 months, 1 year)	Group-based trajectory modelling	Not mentioned	Based on previous study in adults	Two: Early recovery, late recovery
Ashrafioun, 2022, USA [45]	Veterans using primary data from observational study	747 participants; Age: 31.8; 10.3% female	Four measures of 11-point NRS	5 times over 12 months (baseline, 1,3,6,12 months)	Growth mixture modelling	Not mentioned.	BIC, AIC, LMR-LRT, BLRT, entropy	Three: mild increasing to severe, low decreasing, moderate stable

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Page, 2015, Canada [46]	Chronic post-surgical pain following TKA using primary data of observational study	173 participants; Age: 62.9; 49% female	WOMAC	5 times over 12 months (baseline, 4th day, 6 weeks, 3, 12 months)	Growth mixed models	Not mentioned	AIC, interpretability	Four: high pain and late decrease, low pain and gradual decrease, high pain and gradual decrease, high pain
da Silva, 2022, Brazil [47]	Low back pain with secondary analysis of prospective cohort study	542 participants; Age: 68; 86% female	11-point NRS (into three categories)	5 times over 12 months (every 3 months)	Latent class growth analysis	At least three data points required	AIC, BIC, entropy, cluster size > 5%, posterior probabilities, distinctive course	Three: Pain recovery, incomplete pain recovery, persistent severe pain
Simons, 2018, USA [48]	Pain in children following rehabilitation programme using primary data of observational study	253 participants; Age: 14.5; 84% female	11-point NRS on rest	5 times over 1 year (admission, discharge, 1, 4, 12 months)	Latent class growth modelling	Only included those with ≥ 3 data points	BIC, visual interpretation, posterior probabilities, clinically meaningful	Three: early responders, late responders, non-responders
Knecht, 2020, Switzerland [49]	Mid back pain using primary data of an observational study	90 participants; Age: 37.0; 54.4% female	11-point NRS for present pain intensity	6 times over 12 months (baseline and 1 week and 1, 3, 6, 12 months after treatment)	*k*-means clustering adapted for longitudinal data (*KmL* package)	Removed if ≤ 5/6 data points. Non-parametric imputation	Calinski Harabatz 1, Ray Turi, Davies Bouldin, Calinski Harabatz 2, Calinski Harabatz 3	Two: Improving pain, non-improving pain

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Liu, 2021, Canada [50]	Postoperative pain using primary data from observational study	237 participants; Age: 58.5; 51.1% female	11-point NRS	6 times over 12 months (1, 2, 3 days, 3, 6, 12 months)	Growth mixture modelling	Not mentioned	AIC, BIC, parsimony, interpretability, cluster size >5%, posterior probability >0.7	Three: mild pain, moderate pain decreasing, moderate pain sustained
Macedo, 2014, Australia [51]	Non-specific lower back pain using secondary analysis of randomized control trial	155 participants; Age: 49.3; 60.6% female	11-point NRS of average pain in past week	12 times over 1 year (monthly)	Hierarchical clustering with complete linkage	Removed if missing ≥ 3 scores. Others imputed with mean score.	Amalgamation coefficients, ANOVA with Brown-Forsythe test	Fluctuations: Three: nonfluctuating & two others. Severity: Three: Recovering mild, persistent moderate, severe chronic
Ogollah, 2018, UK [52]	Leg pain for patients with lower back pain using secondary analysis of prospective cohort study	609 participants; Age: 50; 63% female	11-point NRS for least, usual and current pain	12 times over 12 months (monthly)	Growth mixture models	All available data used; Sensitivity analysis conducted	BIC, bootstrapped parametric LRT, posterior probabilities ≥ 0.7, clinical interpretability, class size	Four: improving mild, persistent moderate, persistent severe, improving severe

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Hallman, 2018, Denmark [53]	Neck shoulder pain using secondary analysis of cohort study	748 participants; Age: 44.8; 45% female	11-point NRS for peak pain in past month	14 times over 12 months (4 weekly)	Latent class growth analysis	All available data used	BIC, entropy, BLRT, interpretability	Six: Asymptomatic, very low pain, low recovering, moderate fluctuating, strong fluctuating pain, severe persistent pain
Hallman, 2019, Denmark [54]	Neck shoulder pain in employees using secondary analysis of cohort study	748 participants; Age: 44.8; 45% female	11-point NRS for worst pain in past month	14 times over 12 months (4 weekly)	Latent class growth analysis	All available data used	BIC, entropy, BLRT	Six: Asymptomatic, very low, low recovering, moderate fluctuating, strong fluctuating, severe persistent
Huo, 2023, Australia [55]	Low back pain with secondary analysis of observational study	329 participants; Age (median): 56.6; 73% female	11-point NRS	52 times over 52 weeks (weekly)	Latent class growth analysis	At least ten data points required	AIC, BIC, BLRT, entropy, posterior probabilities > 0.7, class size > 5%, clinically distinct	Two: constant low-severity, severe
Kongsted, 2015, Denmark [56]	Low back pain using primary data of observational study	1082 participants; Age: 44; 47% female	Number of days of pain in past week and 11-point NRS for intensity	52 times over 52 weeks (weekly)	Latent class analysis	Participants required ≥ 26 data points	Group size ≥ 5%, BIC increases < 1%	Twelve models used, finding 5-12 clusters each including recovery, mild, moderate or severe alongside ongoing, episodic, slow improvement, relapse

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Pico-Espinosa, 2019, Sweden [57]	Neck pain using secondary analysis of a randomized control trial	617 participants; Age: 46; 69% female.	11-point NRS measuring worst and average pain in the last month and current pain	52 times over 52 weeks (weekly)	Latent class mixed modelling	Not mentioned	BIC	Six: Slightly fluctuating, small improvement, moderate improvement, large improvement, persistent, largely fluctuating
Oakman, 2022, Australia [58]	Musculoskeletal pain using secondary analysis of observational study	488 participants; Age (median) 36-55; 76% female	5 measures of 5-point ordinal scale	3 times over 13 months	Growth mixture modelling	At least two data points required.	BIC	Four: high stable, mid-decrease, low stable, rapid increase
Johnson, 2021, USA [59]	Knee pain using secondary analysis of observational data	188 participants; Age: 58; 63% female	Three measures of 11-point NRS	7 times over 18 months (Quarterly)	Group-based trajectory modelling	Models fit to all available data.	BIC, Bayes factor ≥10	Four: low, moderate-low, moderate-high, high
Crook, 1996, Canada [60]	Work-related musculoskeletal injury using primary analysis of observational study	148 participants; Age: 40.6; 47% female	Short Form McGill Pain Questionnaire	5 times over 21 months (Baseline, 3, 9, 15, 21 months)	*k*-means clustering	Not mentioned	Three clusters chosen (no clear reason)	Three: High level impaired, mid level impaired, low level impaired

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Hebert, 2019, Canada [61]	Post-operative pain for lumbar spinal stenosis surgery using secondary analysis of observational study	548 participants; Age: 66.7; 45.6% female	11-point NRS for average pain in leg and back over past day	4 times over 2 years (pre-operative, 3, 12, 24 months)	Group based trajectory model	Excluded if missing baseline or ≥ 2 others	BIC, group size ≥ 5%, posterior probabilities ≥ 0.7	Three for each leg pain and back pain: Excellent outcome, good outcome, poor outcome
Hebert, 2020, Canada [62]	Degenerative lumbar spinal stenosis using secondary analysis of observational study	529 participants; Age: 66.5; 45.8%	11-point NRS for leg and low back pain over past 24 hours	4 times over 2 years (preoperative, 3, 12, 24 months)	Group-based multi-trajectory model	Participants required 2 follow-up scores and 3 pre-operative scores	Univariate outcome trajectories, BIC, posterior probabilities ≥ 70%, odds of correct classification > 5, clinical judgement	Three: excellent outcome, good outcome, poor outcome
Muller, 2020, UK [63]	Polymyalgia rheumatica, secondary analysis of cohort study	652 participants; Age: 72.4; 62.3% female	11-point NRS	7 times over 2 years (Baseline, 1, 4, 8, 12, 18, 24 months)	Latent class growth analysis	Models fit to all available data, sensitivity analyses to test assumptions	AIC, BIC, sample-size adjusted BIC, boostrapped LRT, VLMR LRT, LMR LRT	Five: classical symptoms, sustained symptoms, partial recovery, recovery then worsening, slow recovery

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Buta, 2022, USA [64]	Back pain in veterans using electronic health record data	117,126 participants; Age: 49.2; 11% female	11-point NRS	24 times over 2 years (monthly)	Latent class growth analysis	Models fit to all available data	BIC, posterior probabilities > 0.7, interpretability	Six: No pain/no opioid, mild pain/no opioid, moderate pain/no opioid, moderate decreasing pain/decreasing opioid, moderate pain/high opioid, moderate increasing pain/increasing opioid
Haybatollahi, 2022, UK [65]	Knee pain using secondary analysis of cohort study	2141 participants; Age: 62.3; 59% female	3 questionnaires totalling 39-point score	3 times over 3 years (Baseline, 1, 3 years)	Latent class growth analysis	At least baseline and one other data point required. Sensitivity analyses conducted.	BIC, bootstrap LRT, VLMR-LRT, entropy, interpretability, theoretical significance	Six for two measures and four for the other. Low stable, moderate recovering, moderate worsening, high stable, worsening, recovering
Enthoven, 2016, Netherlands [66]	Back pain using secondary analysis of prospective cohort study	675 participants; Age: 66.4; 59% female	11-point NRS for average pain over past week	8 times over 3 years (baseline, 6 weeks, 3, 6, 9 months, 1, 2, 3 years)	Latent class growth analysis	Multiple imputation	Indices of fit, usefulness, entropy	Three: low pain, high pain, intermediate pain
Verkleij, 2012, Netherlands [67]	Hip OA using secondary analysis of randomized control trial	222 participants; Age: 63.4; 69.4% female	VAS 0–100mm and WOMAC	9 times over 3 years (3 monthly)	Latent class growth analysis	Not mentioned	BIC, LMR-LRT, BLRT, entropy, usefulness	Five: mild pain, moderate pain, always pain, regularly progressing, highly progressing

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Dunn, 2011, USA [68]	Pain in adolescents using primary data of observational study	1336 participants; Age: 11; 53% female	Presence of pain in four places (yes/no) with follow-up on pain severity and persistence	11 times over 3 years (3 monthly)	Latent class growth analysis	Excluded if missing more than one third of data	AIC, BIC, CAIC, bootstrap *p*-values, size, distinctiveness	Six for back pain, four for facial pain, four for headaches, four for stomach pain
Radojcic, 2020, UK [69]	Knee OA using secondary analysis of a randomized control trial	474 participants; Age: 64.0; 61% female	WOMAC score for pain in indexed knee over past 2 days	7 times over 3.5 years (6 monthly)	Group-based trajectory modelling	Not mentioned	BIC, Wald test, average posterior probabilities	Four: low fluctuating, mild increasing, moderate treatment-sensitive, high treatment-insensitive
Dampier, 2014, USA [70]	Sickle cell disease in children using primary data of observational study	103 participants; Age (median): 7.2 months; 42% female	Presence or absence of pain	At least 60 times over at least 60 days (median 3.8 years; daily)	Latent class growth analysis using pain episodes as proportion per year	Proportions calculated for any year with observations	BIC	Four: no/few pain, low pain, early/intermediate high pain, increasing high pain
Gunzler, 2022, USA [71]	Parkinsons Disease with secondary analysis of observational study	16863 participants; Age: 65.7; 46% female	5-point ordinal scale	9 times over 4 years (every 6 months)	Latent class growth analysis	Models fit to all available data	Entropy, AIC, BIC, aBIC, AICC, cluster size	Five for each disease duration with different descriptions

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Coenen, 2017, Australia [72]	Lower back pain using secondary analysis of cohort study	1249 participants; Age: 17; 53% female	Nordic musculoskeletal pain questionnaire	3 times over 5 years (Age 17, 20, 22)	Latent class analysis for binary outcomes	Only included if ≥ 2 time points	AIC, BIC, posterior probability, entropy, R^2, parsimony	Four: low, increasing, decreasing, high
Sieberg, 2013, USA [73]	Pain in adolescents using secondary analysis of prospective study	190 participants; Age: 14; 72% female	Scoliosis Research Society-30 measure	4 times over 5 years (Preoperatively, 1, 2, 5 years)	Latent class growth modelling	All available data used	BIC, inspection, parsimony, clinical interpretability	Five: no pain, pain improvement, short-term pain, delayed pain, high pain
Arnison, 2022, Sweden [74]	Chronic pain in adolescents using secondary analysis of observational study	2755 participants; Age: 13.7; 47.6% female	10-point NRS	5 times over 5 years (yearly)	Growth mixture modelling	Full information maximum likelihood	aBIC, LMR LRT, entropy	Four: low pain and insomnia, persistent high symptoms, increasing, decreasing
Bastick, 2016, Netherlands [75]	Hip OA using secondary data: a subgroup of an observational study	545 participants; Age: 55.7; 81% female	11-point NRS of pain over the last week	5 times over 5 years (annually)	Latent class growth analysis	Removed if missing ≥ 2 data points	BIC, LMR-LRT, BLRT, entropy, group size (≥ 15%), clinical interpretability	Four: constant mild pain, moderate pain with pain regression, moderate pain with pain progression, constant severe pain
Dainty, 2021, UK [76]	Knee or hip pain using secondary analysis of registry	30,295 participants; Age: 68.8; 58.1% female	Questionnaire with 25-point or 29-point score for pain in hip or knee	5 times over 5 years (baseline, 6 months, 1,3,5 years)	*k*-means clustering	At least four data points required	Non-parametric tests, AIC, BIC	Two: Level 1 (sustained improvement) and Level 2 (sustained improvement at a lower level)

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Dowsey, 2015, Australia [77]	Knee pain post-TKA using secondary analysis of data from a registry	689 participants; Age: 70.5; 67.6% female	7-point scale: Knee Scoring System	5 times over 5 years (annually)	Latent class growth analysis	Not mentioned	AIC, BIC, CAIC, entropy, R^2, convergence, posterior probabilities	Three: no pain, mild pain, moderate pain
Eriks-Hoogland, 2014, Netherlands [78]	Shoulder pain in people with spinal cord injury using secondary analysis of observational study	225 participants; Age: 40.7; 25.3% female	6-point NRS for both shoulders	5 times over 5 years (when able to sit in wheelchair for 4 hours, 3 months, discharge, 1, 5 years)	Latent class growth mixture modelling	Expectation maximisation algorithm	BIC, BLRT, posterior probabilities, clinical relevance	Three: no/low pain, high pain, decreasing pain
Song, 2020, USA [79]	Traumatic brain injury using secondary analysis of observational study	43697 participants; Age: 31.1; 8.4% female	11-point NRS (past year)	5 times over 5 years (annually)	Latent class growth mixed models	Only included if at least one time point per year	AIC, BIC, aBIC, VLMR-LRT, clinical logic	Five: based on treatment, impact and stability of pain
Wieczorek, 2020, France [80]	Hip or knee OA using secondary analysis of observational study	807 participants; Age: 61.9; 68.9% female	Medical Outcomes Study Short Form 36 Questionnaire	5 times over 5 years (annually)	Group-based multi-trajectory modelling	Included if ≥ 2 data points. Maximum likelihood estimation used.	BIC, group size ≥ 5%, posterior probabilities ≥ 0.7, entropy, Cramer's V statistic	Four: severe, moderate, mild, no pain

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Bastick, 2016, Netherlands [81]	Knee OA using secondary analysis of observational study	705 participants; Age: 56.0; 81% female	11-point NRS	6 times over 5 years (baseline and annually)	Latent class growth analysis	Excluded if > 2 data points missing	BIC, entropy, BLRT, VLMR-LRT, clinical relevancy, interpretability	Six: constant mild, moderate regression, major regression, severe progression, moderate progression, constant severe
Wesseling, 2015, Netherlands [82]	Knee OA using secondary analysis of observational study	705 participants; Age: 56; 81% female	11-point NRS of pain in previous week	6 times over 5 years (baseline and annually)	Latent class growth analysis	Only included if ≥ 3 data points	BIC, entropy, BLRT, VLMR-LRT, research question, parsimony, group size, interpretability	Three: marginal, mild, moderate
Shiff, 2018, Canada [83]	Early diagnosis of Juvenile Idiopathic Arthritis using secondary analysis of observational study	1062 participants; Age: 8.8; 63.8% female	VAS 0-10cm	8 times over 5 years (Baseline, 6, 12, 18, 24 months then annually)	Latent trajectory analysis	All available data used	BIC, posterior probabilities approaching 70%, clinical relevance, group size ≥ 5%	Five: minimal, mild-decreasing, mild-increasing, moderate-decreasing, chronically moderate
Cole, 2023, USA [84]	Pain presence in nursing home residents, secondary analysis of data	4864 participants; Age: 80; 63.9% female	Pain presence (binary outcome)	Maximum 28 reports over 5 years	Group-based trajectory modelling	Models fit to all available data	BIC, posterior probabilities > 0.7, odds of correct classification ≥ 5	Four: Consistent pain absence, decreasing-increasing pain presence, increasing-decreasing pain presence, persistent pain presence

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Nicholls, 2014, UK [85]	Knee OA using secondary analysis of observational study	570 participants; Age: 64; 54% female	WOMAC	4 times over 6 years (18, 36, 54, 72 months)	Latent class growth analysis	All available data used	AIC, BIC, ABIC, VLMR-LRT, BLRT, entropy, group size ≥ 1%, posterior probabilities, model interpretability	Five: mild and non-progressive, progressive, moderate, improving, severe and non-improving
Battaglia, 2020, Canada [86]	Pain in adolescents using secondary analysis of twin study	929 participants; Age: 12; 52% female	Sum of 6 4-point NRS	5 times over 6 years (Age 12, 13, 14, 15, 17)	Growth mixture model	Only included if ≥ 2 time points	BIC, AIC, theoretical likelihood index, VLMR-LRT, group size ≥ 5%	Three: none-to-minimal, sporadic, frequent
Collins, 2014, USA [87]	Knee OA using secondary analysis of observational study	1753 participants; Age: 62; 59% female	WOMAC	7 times over 6 years (Baseline and annually)	Growth based trajectory modelling	Acknowledgment of more work required	BIC, group size ≥ 50, stable over different orders	Five: none, mild, low moderate, high moderate, severe
Beynon, 2022, Australia [88]	Spinal pain in adolescents using secondary analysis of observational data	916 participants; Age: 8.4; 52% female	Number of weeks reporting presence of pain	11 times over 6 years (every 6 months)	Latent class growth analysis	Multiple imputation	BIC, clinical judgement, posterior probabilities ≥ 70%, minimum odds ≥ 5	Five: no pain, rare, rare increasing, moderate increasing, early-onset decreasing

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Hebert, 2022, Denmark [89]	Spinal pain in children using secondary analysis of cohort study	1556 participants; Age: 9.1; 52.4% female	Number of weeks reporting presence of pain	11 times over 6 years (every 6 months)	Latent class growth analysis	Imputed with random hot deck multiple imputations	BIC, clinically meaningful, posterior probability ≥ 0.7, odds of correct classification > 5	Five: no pain, rare, rare increasing, moderate increasing, early-onset decreasing
Paquet, 2019, Canada [90]	Vulvodynia using primary data of observational study	173 participants; Age: 31; 100% female	11-point NRS over past six months	3 times over 7 years (Baseline, 2, 7 years)	Latent class growth analysis	Full information maximum likelihood function	BIC, LMR-LRT, BLRT, entropy	Two: Persistent pain, pain reduction
Marshall, 2023, UK [91]	Foot pain using primary data from observational study	425 participants; Age: 64.4; 52.7% female	11-point NRS	5 times over 7 years (baseline, 18, 36, 54, 84 months)	Latent class growth analysis	Baseline and two other data points required. Sensitivity analyses conducted.	AIC, BIC, aBIC, LMR LRT, BLRT, entropy, posterior probabilities >0.7, class size >= 4%, clinical relevance, interpretability	Four: mild improving, moderate improving, moderate-severe persistent, severe persistent
Lee, 2018, South Korea [92]	Pain in older people using secondary analysis of longitudinal study	2839 participants; Age: 68.6; 58% female	Sum of binary variables (pain presence)	5 times over 8 years (biannually)	Group based trajectory modelling	Only complete data used	BIC, parsimony	Three for each gender: low, medium, high

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Dai, 2017, USA [93]	Knee OA using secondary analysis of prospective cohort study	8940 participants; Age: 61.2; 57.9% female	WOMAC	9 times over 8 years (Baseline and annually)	Group-based trajectory model	Not mentioned	BIC, posterior probabilities	Four: none, mild, moderate, severe
Beynon, 2021, Australia [94]	Low back pain using secondary analysis of cohort study	1513 participants; Age: 14; Sex not reported	Presence of pain (binary variable)	4 times over 9 years (age 14, 17, 20, 22)	Group-based trajectory modelling	At least two data points required.	BIC, class size > 30, posterior probabilities > 0.7, odds of correct classification > 5, meaningful distinction	Three: consistently low, decreasing, increasing
Radojcic, 2020, USA [69]	Knee OA using secondary analysis of an observational trial	4796 participants; Age: 61.3; 58.5% female	WOMAC score for pain in the indexed knee over the previous week	10 times over 10 years (annually)	Group-based trajectory model	Not mentioned	BIC, Wald test, average posterior probabilities	Six: Minimal to neglected pain, low fluctuating, mild increasing, moderate treatment-sensitive (cubic trajectory), moderate treatment-sensitive (quadratic trajectory), high treatment-insensitive
Neupane, 2020, Finland [95]	General pain using primary data of observational study	6572 participants; Age: 49.63; 82% female	5-point NRS on various body sites for ongoing pain	3 times over 10 years (5 yearly)	Latent class growth analysis	Not mentioned	AIC, BIC, adjusted BIC, entropy, posterior probabilities, interpretability	Four: high, low, increasing, decreasing

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Pan, 2018, Australia [96]	Knee pain using secondary analysis of observational study	963 participants; Age: 62.8; 50% female	WOMAC score for pain	4 times over 10.7 years (average at 2.6, 5.2, 10.7 years after baseline)	Group-based trajectory model	Sensitivity analysis based on previous time points	AIC, BIC, group size ≥ 10%, posterior probabilities ≥ 0.7	Three: Minimal pain, mild pain, moderate pain
Pan, 2020, Australia [97]	Knee pain using secondary analysis of an observational study	985 participants; Age: 62.9; 50% female	WOMAC score for pain	4 over 10.7 years (average at 2.6, 5.1, 10.7 years after baseline)	Group based trajectory models	Not mentioned	AIC, BIC, group size (≥ 10%), clinical interpretability, posterior probabilities	Three: Minimal pain, mild pain, moderate pain
Airili, 2014, Finland [98]	Musculoskeletal pain in firefighters using primary data of an observational study	360 participants; Age: 48.7; 0% female	Estimated number of days in pain over 12 months	3 times over 13 years (Baseline, 3, 13 years)	Latent class growth modelling	Missing data was assumed as zero pain	AIC, BIC, group size (≥ 5%), clinical interpretability, posterior probabilities	Three: Low symptoms, high pain, high depression
James, 2018, UK [99]	General (with sub analysis for arthritis and cancer) using secondary analysis of an observational study	11977 participants; Age: 63.84; 55.93% female	Binary with pain presence and 3-point severity if present	7 times over 14 years (biannually)	Latent class growth models	Latent class dropout model	BIC, entropy, group size (≥ 5%), clinical interpretability	Four: low or no chronic pain with mild progress, increasing chronic pain, decreasing chronic pain, severe chronic pain with a mild regress

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Mun, 2021, USA [100]	Paediatric pain with secondary analysis of observational study	731 participants; Age: 2 (at baseline); 49% female	3 measures with 3-point ordinal scale	10 times over 14 years (baseline, 1, 2, 3, 5.5, 6.5, 7.5, 8.5, 12, 14 years)	Growth mixture modelling	Full information maximum likelihood	BIC, aBIC, VLMR LRT, BLRT, entropy, parsimony, interpretability, class size	Three: low pain, increasing pain, u-shaped pain
Kyronlahti, 2022, Finland [101]	Low back pain with secondary analysis of observational study	6257 participants; Age: 49.6; 63% female	3 responses on ordinal scale	4 times over 16 years (baseline, 4, 11, 16 years)	Latent class growth analysis	At least two data points required	AIC, BIC, aBIC, entropy, posterior probabilities, interpretability	Three: high-decreasing, intermediate-stable, low
Canizares, 2019, Canada [102]	Back pain using secondary analysis of an observational study	12782 participants; Age: median 39; 51% female	Existence of chronic back pain diagnosed by a professional lasting for six months	9 times over 16 years (biannually)	Group-based trajectory modelling	Assumed missing at random. Missing data imputed using maximum likelihood estimators.	BIC, group size (≥ 5%), posterior probabilities	Four: persistent, developing, recovering, occasional
Milani, 2022, Mexico [103]	Chronic pain using secondary analysis of cohort study	9824 participants; Age: 61.7; 56.2% female	Binary outcome "yes/no" of pain presence	5 times over 17 years	Latent class mixture models	Dropout modelled as function of previous observations	BIC, posterior probabilities > 0.7, class size > 10%	Two: low stable, moderate increasing

First author/ Year/ Country	Pain condition and data source investigated	Participants: number, mean age, % female	Pain score	Frequency of pain score collection	Statistical method used for clustering	Method(s) for dealing with missing data	Criteria for comparison of models	Number and description of clusters
Wickrama, 2021, UK [104]	Physical pain using secondary analysis of observational study	244 participants; Age: 44; 50% female	Two measures of 6-point and 5-point NRS	3 times over 21 years (baseline, 7, 21 years)	Latent class trajectory analysis	Complete data only	aBIC, entropy, interpretability, cluster size >5%, LMR-LRT	Three for wives and three for husbands: high decreasing, moderate increasing, low
Aili, 2021, Sweden [105]	Chronic regional pain & Chronic widespread pain using primary data of an observational study	1858 participants; Age: 46.3; 55.3% female	Presence of pain in 3/12 previous months. Pain area reported on a manakin	5 times over 21 years (3, 8, 12, 21 years after baseline)	Latent class growth analysis	Individuals removed if missing two or more data points	AIC, BIC, LMR-LRT, group size (≥ 10%), clinical interpretability, posterior probabilities	Five: persistent no chronic pain, persistent chronic widespread pain, persistent chronic regional pain, migration from chronic regional pain to chronic widespread pain, change from no chronic pain to chronic regional pain or chronic widespread pain
Leino-Arjas, 2018, Sweden [106]	Musculoskeletal pain using secondary analysis of cohort study	1066 participants; Age: 16; 47.2% female	3-point scale for pain in various body areas	4 times over 27 years (Age 16, 21, 30, 43)	Latent class growth analysis	All available information used	BIC, LMR-LRT, entropy, group size, posterior probabilities, interpretability	Men: Three: High-stable, low-increasing, moderate-stable. Women: Three: high-stable, low-increasing, moderate-decreasing
Neupane, 2018, Finland [107]	Multisite musculoskeletal pain using secondary analysis of longitudinal study	3093 participants; Age: 49.7; 62.8% female	3-point scale for pain in various body areas	5 times over 28 years (Baseline, 4, 11, 16, 28 years)	Growth mixed models	Not mentioned	AIC, BIC, adjusted BIC, entropy, LMR-LRT, meaningfulness, interpretability	Three: low, moderate, high-decreasing

Abbreviations: aBIC: Sample size adjusted Bayesian Information Criteria, AIC: Akaike Information Criteria, AICC: Akaike Information Criteria with correction, ANOVA: Analysis of Variance, BIC: Bayesian Information Criteria, BLRT: Bootstrap Likelihood Ratio Test, CAIC: Consistent Akaike Information Criteria, KOOS: Knee injury and Osteoarthritis Outcome Score, LMR LRT: Lo-Mendell-Rubin Likelihood Ratio Test, LRT: Likelihood Ratio Test, NRS: Numeric Rating Scale, OA: Osteoarthritis, PCA: Principal Component Analysis, RMSEA: Root Mean Square Error of Approximation, TKA: Total Knee Athroplasty, VLMR LRT: Vuong-Lo-Medell-Rubin Likelihood Ratio Test, WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index

References

1. Hundert A, Woolcott C, Dorling J, et al (2021) Classification of Individual Pain Response Trajectories Following Medically Indicated Heel Lances in Preterm Infants During Their NICU Admission. Clin J Pain 38:151–158. https://doi.org/10.1097/AJP.0000000000001011

2. Kannampallil T, Galanter WL, Falck S, et al (2016) Characterizing the pain score trajectories of hospitalized adult medical and surgical patients: a retrospective cohort study. Pain 157:2739–2746. https://doi.org/10.1097/j.pain.0000000000000693

3. Weng Y, Tian L, Tedesco D, et al (2020) Trajectory analysis for postoperative pain using electronic health records: A nonparametric method with robust linear regression and K-medians cluster analysis. Health Informatics J 26:1404–1418. https://doi.org/10.1177/1460458219881339

4. Pagé MG, Karanicolas PJ, Cleary S, et al (2019) In-hospital opioid consumption, but not pain intensity scores, predicts 6-month levels of pain catastrophizing following hepatic resection: A trajectory analysis. Eur J Pain Lond Engl 23:503–514. https://doi.org/10.1002/ejp.1324

5. Wu M-C, Lu T-C, Cheng M-T, et al (2022) Pain trajectories in the emergency department: Patient characteristics and clinical outcomes. Am J Emerg Med 55:111–116. https://doi.org/10.1016/j.ajem.2021.09.087

6. Thomazeau J, Rouquette A, Martinez V, et al (2016) Predictive Factors of Chronic Post-Surgical Pain at 6 Months Following Knee Replacement: Influence of Postoperative Pain Trajectory and Genetics. Pain Physician 19:E729-741

7. Awadalla SS, Winslow V, Avidan MS, et al (2022) Effect of acute postsurgical pain trajectories on 30-day and 1-year pain. PloS One 17:e0269455. https://doi.org/10.1371/journal.pone.0269455

8. Althaus A, Arránz Becker O, Moser K-H, et al (2018) Postoperative Pain Trajectories and Pain Chronification—an Empirical Typology of Pain Patients. Pain Med 19:2536–2545. https://doi.org/10.1093/pm/pny099

9. Getachew M, Lerdal A, Småstuen MC, et al (2023) Modifiable factors associated with a consistently high acute pain trajectory after surgical treatment of traumatic fractures in Ethiopia: a multi-center prospective cohort study. J Orthop Surg 18:288. https://doi.org/10.1186/s13018-023-03770-0

10. Ilhan E, Galea C, Pacey V, et al (2020) Trajectories of post-surgical pain in infants admitted to neonatal intensive care. Eur J Pain Lond Engl 24:1822–1830. https://doi.org/10.1002/ejp.1629

11. Pagé MG, Katz J, Curtis K, et al (2016) Acute pain trajectories and the persistence of post-surgical pain: a longitudinal study after total hip arthroplasty. J Anesth 30:568–577. https://doi.org/10.1007/s00540-016-2183-4

12. Ocay DD, Li MMJ, Ingelmo P, et al (2020) Predicting Acute Postoperative Pain Trajectories and Long-Term Outcomes of Adolescents after Spinal Fusion Surgery. Pain Res Manag 2020:9874739. https://doi.org/10.1155/2020/9874739

13. Thomas MJ, Yu D, Nicholls E, et al (2020) Short-Term Recovery Trajectories of Acute Flares in Knee Pain: A UK-Netherlands Multicenter Prospective Cohort Analysis. Arthritis Care Res 72:1687–1692. https://doi.org/10.1002/acr.24088

14. Mamoun N, Wright MC, Bottiger B, et al (2022) Pain Trajectories After Valve Surgeries Performed via Midline Sternotomy Versus Mini-Thoracotomy. J Cardiothorac Vasc Anesth 36:3596–3602. https://doi.org/10.1053/j.jvca.2022.05.007

15. Chang W-K, Li Y-S, Wu H-L, et al (2022) Group-based trajectory analysis of postoperative pain in epidural analgesia for video-assisted thoracoscopic surgery and risk factors of rebound pain. J Chin Med Assoc JCMA 85:216–221. https://doi.org/10.1097/JCMA.0000000000000647

16. Vasilopoulos T, Wardhan R, Rashidi P, et al (2021) Patient and Procedural Determinants of Postoperative Pain Trajectories. Anesthesiology 134:421–434. https://doi.org/10.1097/ALN.0000000000003681

17. Jaffa MN, Jha RM, Elmer J, et al (2021) Pain Trajectories Following Subarachnoid Hemorrhage are Associated with Continued Opioid Use at Outpatient Follow-up. Neurocrit Care 35:806–814. https://doi.org/10.1007/s12028-021-01282-5

18. Daoust R, Paquet J, Cournoyer A, et al (2019) Acute Pain Resolution After an Emergency Department Visit: A 14-Day Trajectory Analysis. Ann Emerg Med 74:224–232. https://doi.org/10.1016/j.annemergmed.2019.01.019

19. Cummings C, Ewing SWF, Tran DD, et al (2022) Trajectories and biopsychosocial predictors of daily acute pain in adolescents receiving treatment for pain: a daily diary study. J Behav Med 45:613–621. https://doi.org/10.1007/s10865-022-00297-3

20. Vowles KE, Witkiewitz K, Levell J, et al (2017) Are reductions in pain intensity and pain-related distress necessary? An analysis of within-treatment change trajectories in relation to improved functioning following interdisciplinary acceptance and commitment therapy for adults with chronic pain. J Consult Clin Psychol 85:87–98. https://doi.org/10.1037/ccp0000159

21. Walton DM, Eilon-Avigdor Y, Wonderham M, Wilk P (2014) Exploring the clinical course of neck pain in physical therapy: a longitudinal study. Arch Phys Med Rehabil 95:303–308. https://doi.org/10.1016/j.apmr.2013.09.004

22. Mori M, Brooks C, Dhruva SS, et al (2021) Trajectories of Pain After Cardiac Surgery: Implications for Measurement, Reporting, and Individualized Treatment. Circ Cardiovasc Qual Outcomes 14:e007781. https://doi.org/10.1161/CIRCOUTCOMES.120.007781

23. Connor JP, Brier ZMF, Price M (2020) The Association Between Pain Trajectories With Posttraumatic Stress Disorder, Depression, and Disability During the Acute Posttrauma Period. Psychosom Med 82:862–868. https://doi.org/10.1097/PSY.0000000000000866

24. Shearer HM, Côté P, Hogg-Johnson S, et al (2022) Pain trajectories and well-being in children and young people with cerebral palsy: A cohort study. Dev Med Child Neurol 64:1416–1424. https://doi.org/10.1111/dmcn.15252

25. Singh JA, Lemay CA, Nobel L, et al (2019) Association of Early Postoperative Pain Trajectories With Longer-term Pain Outcome After Primary Total Knee Arthroplasty. JAMA Netw Open 2:e1915105. https://doi.org/10.1001/jamanetworkopen.2019.15105

26. Downie AS, Hancock MJ, Rzewuska M, et al (2016) Trajectories of acute low back pain: a latent class growth analysis. PAIN 157:225–234. https://doi.org/10.1097/j.pain.0000000000000351

27. Lee AC, Harvey WF, Han X, et al (2018) Pain and functional trajectories in symptomatic knee osteoarthritis over up to 12 weeks of exercise exposure. Osteoarthritis Cartilage 26:501–512. https://doi.org/10.1016/j.joca.2018.01.016

28. Daoust R, Emond M, Bergeron E, et al (2013) Risk factors of significant pain syndrome 90 days after minor thoracic injury: trajectory analysis. Acad Emerg Med Off J Soc Acad Emerg Med 20:1139–1145. https://doi.org/10.1111/acem.12248

29. Atukorala I, Downie A, Pathmeswaran A, et al (2022) Short-term pain trajectories in patients with knee osteoarthritis. Int J Rheum Dis 25:281–294. https://doi.org/10.1111/1756-185X.14267

30. Burns JW, Janssen I, Lillis T, et al (2020) The transition from acute to persistent pain: the identification of distinct trajectories among women presenting to an emergency department. Pain 161:2511–2519. https://doi.org/10.1097/j.pain.0000000000001960

31. Braunwalder C, Ehrmann C, Hodel J, et al (2022) Pain Trajectories During Initial Rehabilitation After Spinal Cord Injury: Do Psychosocial Resources and Mental Health Predict Trajectories? Arch Phys Med Rehabil 103:1294–1302. https://doi.org/10.1016/j.apmr.2022.01.149

32. Toyoda H, Takahashi S, Hoshino M, et al (2017) Characterizing the course of back pain after osteoporotic vertebral fracture: a hierarchical cluster analysis of a prospective cohort study. Arch Osteoporos 12:82. https://doi.org/10.1007/s11657-017-0377-5

33. Schuller W, Ostelo RW, Rohrich DC, et al (2021) Pain trajectories and possible predictors of a favourable course of low back pain in patients consulting musculoskeletal physicians in The Netherlands. Chiropr Man Ther 29:38. https://doi.org/10.1186/s12998-021-00392-3

34. Dunn KM, Jordan K, Croft PR (2006) Characterizing the Course of Low Back Pain: A Latent Class Analysis. Am J Epidemiol 163:754–761. https://doi.org/10.1093/aje/kwj100

35. Axén I, Bodin L, Bergström G, et al (2011) Clustering patients on the basis of their individual course of low back pain over a six month period. BMC Musculoskelet Disord 12:99. https://doi.org/10.1186/1471-2474-12-99

36. Ailliet L, Rubinstein SM, Hoekstra T, et al (2018) Long-term trajectories of patients with neck pain and low back pain presenting to chiropractic care: A latent class growth analysis. Eur J Pain Lond Engl 22:103–113. https://doi.org/10.1002/ejp.1094

37. Tamcan O, Mannion AF, Eisenring C, et al (2010) The course of chronic and recurrent low back pain in the general population. Pain 150:451–457. https://doi.org/10.1016/j.pain.2010.05.019

38. Brennan PL, SooHoo S (2020) Effects of Mental Health Disorders on Nursing Home Residents' Nine-Month Pain Trajectories. Pain Med Malden Mass 21:488–500. https://doi.org/10.1093/pm/pnz177

39. Dumenci L, Perera RA, Keefe FJ, et al (2019) Model-based pain and function outcome trajectory types for patients undergoing knee arthroplasty: a secondary analysis from a randomized clinical trial. Osteoarthritis Cartilage 27:878–884. https://doi.org/10.1016/j.joca.2019.01.004

40. Ellyson AM, Gordon G, Zhou C, Rabbitts JA (2022) Trajectories, Risk Factors, and Impact of Persistent Pain After Major Musculoskeletal Surgery in Adolescents: A Replication Study. J Pain 23:995–1005. https://doi.org/10.1016/j.jpain.2021.12.009

41. Lee JY, Walton DM, Tremblay P, et al (2020) Defining pain and interference recovery trajectories after acute non-catastrophic musculoskeletal trauma through growth mixture modeling. BMC Musculoskelet Disord 21:615. https://doi.org/10.1186/s12891-020-03621-7

42. Müller M, Bütikofer L, Andersen OK, et al (2021) Cold pain hypersensitivity predicts trajectories of pain and disability after low back surgery: a prospective cohort study. Pain 162:184–194. https://doi.org/10.1097/j.pain.0000000000002006

43. Panken G, Hoekstra T, Verhagen A, et al (2016) Predicting chronic low-back pain based on pain trajectories in patients in an occupational setting: an exploratory analysis. Scand J Work Environ Health 42:520–527. https://doi.org/10.5271/sjweh.3584

44. Rabbitts JA, Zhou C, Groenewald CB, et al (2015) Trajectories of postsurgical pain in children: risk factors and impact of late pain recovery on long-term health outcomes after major surgery. Pain 156:2383–2389. https://doi.org/10.1097/j.pain.0000000000000281

45. Ashrafioun L, Saulnier KG, Allan NP, et al (2022) Pain intensity trajectories among veterans seeking mental health treatment: Association with mental health symptoms and suicidal thoughts and behaviors. J Affect Disord 297:586–592. https://doi.org/10.1016/j.jad.2021.10.082

46. Pagé MG, Katz J, Romero Escobar EM, et al (2015) Distinguishing problematic from nonproblematic postsurgical pain: a pain trajectory analysis after total knee arthroplasty. Pain 156:460–468. https://doi.org/10.1097/01.j.pain.0000460327.10515.2d

47. Silva JP da, de Jesus-Moraleida FR, Felício DC, et al (2022) Trajectories of pain and disability in older adults with acute low back pain: Longitudinal data of the BACE-Brazil cohort. Braz J Phys Ther 26:100386. https://doi.org/10.1016/j.bjpt.2021.100386

48. Simons LE, Sieberg CB, Conroy C, et al (2018) Children With Chronic Pain: Response Trajectories After Intensive Pain Rehabilitation Treatment. J Pain 19:207–218. https://doi.org/10.1016/j.jpain.2017.10.005

49. Knecht C, Hartnack S, Sick B, et al (2020) A prospective observational study on trajectories and prognostic factors of mid back pain. BMC Musculoskelet Disord 21:554. https://doi.org/10.1186/s12891-020-03534-5

50. Liu CW, Page MG, Weinrib A, et al (2021) Predictors of one year chronic post-surgical pain trajectories following thoracic surgery. J Anesth 35:505–514. https://doi.org/10.1007/s00540-021-02943-7

51. Macedo LG, Maher CG, Latimer J, et al (2014) Nature and determinants of the course of chronic low back pain over a 12-month period: a cluster analysis. Phys Ther 94:210–221. https://doi.org/10.2522/ptj.20120416

52. Ogollah RO, Konstantinou K, Stynes S, Dunn KM (2018) Determining One-Year Trajectories of Low-Back-Related Leg Pain in Primary Care Patients: Growth Mixture Modeling of a Prospective Cohort Study. Arthritis Care Res 70:1840–1848. https://doi.org/10.1002/acr.23556

53. Hallman DM, Rasmussen CDN, Jørgensen MB, Holtermann A (2018) Time course of neck-shoulder pain among workers: A longitudinal latent class growth analysis. Scand J Work Environ Health 44:47–57. https://doi.org/10.5271/sjweh.3690

54. Hallman DM, Holtermann A, Dencker-Larsen S, et al (2019) Are trajectories of neck-shoulder pain associated with sick leave and work ability in workers? A 1-year prospective study. BMJ Open 9:e022006. https://doi.org/10.1136/bmjopen-2018-022006

55. Huo M, Ho E, Kongsted A, et al (2023) Association between physical activity, sedentary behaviour and the trajectory of low back pain. Spine J Off J North Am Spine Soc 23:1037–1044. https://doi.org/10.1016/j.spinee.2023.03.006

56. Kongsted A, Kent P, Hestbaek L, Vach W (2015) Patients with low back pain had distinct clinical course patterns that were typically neither complete recovery nor constant pain. A latent class analysis of longitudinal data. Spine J Off J North Am Spine Soc 15:885–894. https://doi.org/10.1016/j.spinee.2015.02.012

57. Pico-Espinosa OJ, Côté P, Hogg-Johnson S, et al (2019) Trajectories of Pain Intensity Over 1 Year in Adults With Disabling Subacute or Chronic Neck Pain: Clin J Pain 35:678–685. https://doi.org/10.1097/AJP.0000000000000727

58. Oakman J, Neupane S, Kyrönlahti S, et al (2022) Musculoskeletal pain trajectories of employees working from home during the COVID-19 pandemic. Int Arch Occup Environ Health 95:1891–1901. https://doi.org/10.1007/s00420-022-01885-1

59. Johnson AJ, Vasilopoulos T, Booker SQ, et al (2021) Knee pain trajectories over 18 months in non-Hispanic Black and non-Hispanic White adults with or at risk for knee osteoarthritis. BMC Musculoskelet Disord 22:415. https://doi.org/10.1186/s12891-021-04284-8

60. Crook J, Moldofsky H (1996) The clinical course of musculoskeletal pain in empirically derived groupings of injured workers. Pain 67:427–433. https://doi.org/10.1016/0304-3959(96)03151-x

61. Hebert JJ, Abraham E, Wedderkopp N, et al (2019) Patients undergoing surgery for lumbar spinal stenosis experience unique courses of pain and disability: A group-based trajectory analysis. PloS One 14:e0224200. https://doi.org/10.1371/journal.pone.0224200

62. Hébert JJ, Abraham E, Wedderkopp N, et al (2020) Preoperative Factors Predict Postoperative Trajectories of Pain and Disability Following Surgery for Degenerative Lumbar Spinal Stenosis. Spine 45:E1421–E1430. https://doi.org/10.1097/BRS.0000000000003587

63. Muller S, Whittle R, Hider SL, et al (2020) Longitudinal clusters of pain and stiffness in polymyalgia rheumatica: 2-year results from the PMR Cohort Study. Rheumatol Oxf Engl 59:1906–1915. https://doi.org/10.1093/rheumatology/kez533

64. Buta E, Gordon KS, Gueorguieva R, et al (2022) Joint longitudinal trajectories of pain intensity and opioid prescription in veterans with back pain. Pharmacoepidemiol Drug Saf 31:1262–1271. https://doi.org/10.1002/pds.5531

65. Haybatollahi SM, James RJE, Fernandes G, et al (2022) Identifying multiple knee pain trajectories and the prediction of opioid and NSAID medication used: A latent class growth approach. Pain Pract Off J World Inst Pain 22:210–221. https://doi.org/10.1111/papr.13082

66. Enthoven WTM, Koes BW, Bierma-Zeinstra SMA, et al (2016) Defining trajectories in older adults with back pain presenting in general practice. Age Ageing 45:878–883. https://doi.org/10.1093/ageing/afw127

67. Verkleij SPJ, Hoekstra T, Rozendaal RM, et al (2012) Defining discriminative pain trajectories in hip osteoarthritis over a 2-year time period. Ann Rheum Dis 71:1517–1523. https://doi.org/10.1136/annrheumdis-2011-200687

68. Dunn KM, Jordan KP, Mancl L, et al (2011) Trajectories of pain in adolescents: A prospective cohort study. Pain 152:66–73. https://doi.org/10.1016/j.pain.2010.09.006

69. Radojčić MR, Arden NK, Yang X, et al (2020) Pain trajectory defines knee osteoarthritis subgroups: a prospective observational study. Pain 161:2841–2851. https://doi.org/10.1097/j.pain.0000000000001975

70. Dampier C, Ely B, Brodecki D, et al (2014) Pain characteristics and age-related pain trajectories in infants and young children with sickle cell disease: Pain Trajectories in Infants With SCD. Pediatr Blood Cancer 61:291–296. https://doi.org/10.1002/pbc.24796

71. Gunzler DD, Gunzler SA, Briggs FBS (2022) Heterogeneous pain trajectories in persons with Parkinson's disease. Parkinsonism Relat Disord 102:42–50. https://doi.org/10.1016/j.parkreldis.2022.07.006

72. Coenen P, Smith A, Paananen M, et al (2017) Trajectories of Low Back Pain From Adolescence to Young Adulthood. Arthritis Care Res 69:403–412. https://doi.org/10.1002/acr.22949

73. Sieberg CB, Simons LE, Edelstein MR, et al (2013) Pain prevalence and trajectories following pediatric spinal fusion surgery. J Pain 14:1694–1702. https://doi.org/10.1016/j.jpain.2013.09.005

74. Arnison T, Schrooten MGS, Bauducco S, et al (2022) Sleep phase and pre-sleep arousal predicted co-developmental trajectories of pain and insomnia within adolescence. Sci Rep 12:4480. https://doi.org/10.1038/s41598-022-08207-y

75. Bastick AN, Verkleij SPJ, Damen J, et al (2016) Defining hip pain trajectories in early symptomatic hip osteoarthritis – 5 year results from a nationwide prospective cohort study (CHECK). Osteoarthritis Cartilage 24:768–775. https://doi.org/10.1016/j.joca.2015.11.023

76. Dainty JR, Smith TO, Clark EM, et al (2021) Trajectories of pain and function in the first five years after total hip and knee arthroplasty : an analysis of patient reported outcome data from the National Joint Registry. Bone Jt J 103-B:1111–1118. https://doi.org/10.1302/0301-620X.103B6.BJJ-2020-1437.R1

77. Dowsey MM, Smith AJ, Choong PFM (2015) Latent Class Growth Analysis predicts long term pain and function trajectories in total knee arthroplasty: a study of 689 patients. Osteoarthritis Cartilage 23:2141–2149. https://doi.org/10.1016/j.joca.2015.07.005

78. Eriks-Hoogland IE, Hoekstra T, de Groot S, et al (2014) Trajectories of musculoskeletal shoulder pain after spinal cord injury: Identification and predictors. J Spinal Cord Med 37:288–298. https://doi.org/10.1179/2045772313Y.0000000168

79. Song K, Wang C-P, McGeary DD, et al (2020) Five-year Pain Intensity and Treatment Trajectories of Post-9/11 Veterans With Mild Traumatic Brain Injury. J Pain 21:1005–1017. https://doi.org/10.1016/j.jpain.2019.12.009

80. Wieczorek M, Rotonda C, Coste J, et al (2020) Trajectory analysis combining pain and physical function in individuals with knee and hip osteoarthritis: results from the French KHOALA cohort. Rheumatol Oxf Engl 59:3488–3498. https://doi.org/10.1093/rheumatology/keaa148

81. Bastick AN, Wesseling J, Damen J, et al (2016) Defining knee pain trajectories in early symptomatic knee osteoarthritis in primary care: 5-year results from a nationwide prospective cohort study (CHECK). Br J Gen Pract J R Coll Gen Pract 66:e32-39. https://doi.org/10.3399/bjgp15X688129

82. Wesseling J, Bastick AN, ten Wolde S, et al (2015) Identifying Trajectories of Pain Severity in Early Symptomatic Knee Osteoarthritis: A 5-year Followup of the Cohort Hip and Cohort Knee (CHECK) Study. J Rheumatol 42:1470–1477. https://doi.org/10.3899/jrheum.141036

83. Shiff NJ, Tupper S, Oen K, et al (2018) Trajectories of pain severity in juvenile idiopathic arthritis: results from the Research in Arthritis in Canadian Children Emphasizing Outcomes cohort. Pain 159:57–66. https://doi.org/10.1097/j.pain.0000000000001064

84. Cole CS, Carpenter JS, Blackburn J, et al (2023) Pain trajectories of nursing home residents. J Am Geriatr Soc 71:1188–1197. https://doi.org/10.1111/jgs.18182

85. Nicholls E, Thomas E, van der Windt DA, et al (2014) Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: findings from the Knee Clinical Assessment Study and the Osteoarthritis Initiative. Osteoarthritis Cartilage 22:2041–2050. https://doi.org/10.1016/j.joca.2014.09.026

86. Battaglia M, Garon-Carrier G, Brendgen M, et al (2020) Trajectories of pain and anxiety in a longitudinal cohort of adolescent twins. Depress Anxiety 37:475–484. https://doi.org/10.1002/da.22992

87. Collins JE, Katz JN, Dervan EE, Losina E (2014) Trajectories and risk profiles of pain in persons with radiographic, symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage 22:622–630. https://doi.org/10.1016/j.joca.2014.03.009

88. Beynon AM, Wedderkopp N, Walker BF, et al (2022) No associations between C-reactive protein and spinal pain trajectories in children and adolescents (CHAMPS study-DK). Sci Rep 12:20001. https://doi.org/10.1038/s41598-022-24587-7

89. Hébert JJ, Beynon AM, Jones BL, et al (2022) Spinal pain in childhood: prevalence, trajectories, and diagnoses in children 6 to 17 years of age. Eur J Pediatr 181:1727–1736. https://doi.org/10.1007/s00431-021-04369-5

90. Pâquet M, Vaillancourt-Morel M-P, Jodouin J-F, et al (2019) Pain Trajectories and Predictors: A 7-Year Longitudinal Study of Women With Vulvodynia. J Sex Med 16:1606–1614. https://doi.org/10.1016/j.jsxm.2019.07.018

91. Marshall M, Blagojevic-Bucknall M, Rathod-Mistry T, et al (2023) Identifying Long-Term Trajectories of Foot Pain Severity and Potential Prognostic Factors: A Population-Based Cohort Study. Arthritis Care Res 75:1123–1131. https://doi.org/10.1002/acr.24823

92. Lee J, Jang S-N, Cho S-I (2018) The Trajectories of the Number of Pain Sites and Their Associated Factors in Older Adults: Results from the Korean Longitudinal Study of Ageing. Gerontology 64:532–540. https://doi.org/10.1159/000490051

93. Dai Z, Lu N, Niu J, et al (2017) Dietary Fiber Intake in Relation to Knee Pain Trajectory. Arthritis Care Res 69:1331–1339. https://doi.org/10.1002/acr.23158

94. Beynon AM, Hebert JJ, Beales DJ, et al (2021) Multi-trajectory analysis of C-reactive protein and low back pain from adolescence to early adulthood. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 30:1028–1034. https://doi.org/10.1007/s00586-020-06677-0

95. Neupane S, Lallukka T, Pietiläinen O, et al (2020) Trajectories of multisite musculoskeletal pain in midlife: Associations with common mental disorders. Eur J Pain Lond Engl 24:364–373. https://doi.org/10.1002/ejp.1492

96. Pan F, Tian J, Aitken D, et al (2018) Predictors of pain severity trajectory in older adults: a 10.7-year follow-up study. Osteoarthritis Cartilage 26:1619–1626. https://doi.org/10.1016/j.joca.2018.08.002

97. Pan F, Tian J, Cicuttini F, Jones G (2020) Metabolic syndrome and trajectory of knee pain in older adults. Osteoarthritis Cartilage 28:45–52. https://doi.org/10.1016/j.joca.2019.05.030

98. Airila A, Hakanen JJ, Luukkonen R, et al (2014) Developmental trajectories of multisite musculoskeletal pain and depressive symptoms: the effects of job demands and resources and individual factors. Psychol Health 29:1421–1441. https://doi.org/10.1080/08870446.2014.945929

99. James RJE, Walsh DA, Ferguson E (2018) General and disease-specific pain trajectories as predictors of social and political outcomes in arthritis and cancer. BMC Med 16:51. https://doi.org/10.1186/s12916-018-1031-9

100. Mun CJ, Lemery-Chalfant K, Wilson M, Shaw DS (2021) Predictors and Consequences of Pediatric Pain Symptom Trajectories: A 14-Year Longitudinal Study. Pain Med Malden Mass 22:2162–2173. https://doi.org/10.1093/pm/pnab173

101. Kyrönlahti SM, Nygård C-H, K C P, Neupane S (2022) Trajectories of low back pain from midlife to retirement and functional ability at old age. Eur J Public Health 32:497–503. https://doi.org/10.1093/eurpub/ckab191

102. Canizares M, Rampersaud YR, Badley EM (2019) Course of Back Pain in the Canadian Population: Trajectories, Predictors, and Outcomes. Arthritis Care Res 71:1660–1670. https://doi.org/10.1002/acr.23811

103. Milani SA, Howrey B, Rodriguez MA, et al (2022) Gender differences in activity-limiting pain trajectories over a 17-year period in the Mexican Health and Aging Study. Pain 163:e285–e292. https://doi.org/10.1097/j.pain.0000000000002292

104. Wickrama KAS, Lee TK, O'Neal CW (2021) Heterogeneity in Husbands' and Wives' Physical Pain Trajectories Over Mid-Later Years: Biopsychosocial Stratification and Implications for Later-Life Well-Being. The Gerontologist 61:930–941. https://doi.org/10.1093/geront/gnab007

105. Aili K, Campbell P, Michaleff ZA, et al (2021) Long-term trajectories of chronic musculoskeletal pain: a 21-year prospective cohort latent class analysis. Pain 162:1511–1520. https://doi.org/10.1097/j.pain.0000000000002137

106. Leino-Arjas P, Rajaleid K, Mekuria G, et al (2018) Trajectories of musculoskeletal pain from adolescence to middle age: the role of early depressive symptoms, a 27-year follow-up of the Northern Swedish Cohort. Pain 159:67–74. https://doi.org/10.1097/j.pain.0000000000001065

107. Neupane S, Nygård C-H, Prakash KC, et al (2018) Multisite musculoskeletal pain trajectories from midlife to old age: a 28-year follow-up of municipal employees. Occup Environ Med 75:863–870. https://doi.org/10.1136/oemed-2018-105235

Appendix C

Appendix to Chapter 2: Table of papers from literature search in Section 2.3

First author, Year	Participants: Number, % female, mean age	Pain Condition	Measure of Pain	Duration of data collection	Frequency of data collection	Measure of variability	Variables associated with variability
Frank, 2013 [1]	14 participants, 42.9% female, Age: 53.1	Spinal cord injury	91-point graphical rating scale	One day	4 times per day	33% difference deemed clinically significant	Not investigated
Allen, 2009 [2]	157 participants, 52% female, Age: 10.6	Osteoarthritis	100-point VAS	2 days	Every 2 hours	Range of pain scores	Within-day variability associated with: BMI, site of pain, medication, age, race
Zempsky, 2013 [3]	25 participants, 80% female, Age: 16.6	Sickle cell disease	11-point NRS	Hospital stay (mean: 6.7 days)	Every 4 hours	Linear mixed models	No significant associations with pain variability
Kratz, 2017 [4]	107 participants, 69% female,	Multiple sclerosis	11-point NRS	7 days	Five times a day	Multilevel modelling	Not investigated

First author, Year	Participants: Number, % female, mean age	Pain Condition	Measure of Pain	Duration of data collection	Frequency of data collection	Measure of variability	Variables associated with variability
	Age: 45.2						
Farrar, 2014 [5]	2740 participants, 48.1% female, Age: 65.8	Clinical trial data	11-point NRS	7 days	Daily	Standard Deviation	Variability associated with response in placebo group
Kuzu, 2022 [6]	168 participants, 37% female, Age: 49.8	Spinal cord injury	PROMIS Pain Intensity Short Form	7 days	Daily	Multivariable model	Days with greater pain intensity associated with: fatigue, depressive symptoms, cognitive function
Schneider, 2012 [7]	194 participants, 78% female, Age: 66	Osteoarthritis	11-point NRS	7 days	Daily	Dispersion model	No significant associations with pain variability
O'Brien, 2011 [8]	22 participants, 100% female, Age: 43.8	Fibromyalgia	VAS	2 weeks	Daily	Hierarchical linear models	Poor sleep associated with higher next-day pain
Erickson, 2021 [9]	20 participants, 55% female, Age: 42.1	Urologic chronic pelvic pain syndrome	11-point VAS	14 days	4 times per day	Standard Deviation	Not investigated

First author, Year	Participants: Number, % female, mean age	Pain Condition	Measure of Pain	Duration of data collection	Frequency of data collection	Measure of variability	Variables associated with variability
Dzierzewski, 2010 [10]	50 participants, 66% female, Age: 69.1	General pain	11-point NRS	14 days	Daily	Multilevel modelling	Days with greater pain associated with time spent asleep
Wesolowicz, 2021 [11]	54 participants, 70.4% female, Age: 36.92	Low back pain	101-cm VAS	14-days	Daily	Standard Deviation	Days with greater pain intensity associated with: Daily catastrophizing, self-efficacy, negative affect
Andersen, 2017 [12]	95 participants, 53% female, Age: 30.9	Low back pain	11-point NRS	3 weeks	Twice daily	Linear mixed models	Days with greater pain associated with occupational load
Martire, 2016 [13]	135 participants, 50% female, Age: 65	Knee OA	4-point NRS	22 days	Three times a day	Multilevel modelling	Genotypes associated with variability. Physical activity associated with days of higher pain severity
Schneider, 2012 [7]	106 participants, 86% female, Age: 56	Rheumatic disease	101-point VAS	28 days	Daily	Dispersion model	Disease (OA had lower variability), depression, happiness, frustration

First author, Year	Participants: Number, % female, mean age	Pain Condition	Measure of Pain	Duration of data collection	Frequency of data collection	Measure of variability	Variables associated with variability
Taylor, 2013 [14]	251 participants, 100% female, Age: 57.3	Fibromyalgia and OA	101-point NRS	30 days	Daily	Deviation from average; multilevel models	Days with greater pain associated with physical disability
Vivekanantham, 2023 [15]	25 participants, 48% female, Age: 65	Knee OA	11-point NRS	90 days	Two times a day	Variance	Low physical activity associated with fluctuating pain
Fogel, 2013 [16]	5 participants, 100% female, Age 31-56	Low back pain	Percent of day in pain	Average 165 days	Daily	Standard Deviation	Not investigated

References

1. Frank AO, Gawronski J, Spyridonis F, Ghinea G (2013) Pain management following new and long-standing spinal cord injury: a pilot study of changes in pain intensity experienced during the day. Int J Rehabil Res Int Z Rehabil Rev Int Rech Readaptation 36:379–382. https://doi.org/10.1097/MRR.0b013e3283646b74

2. Allen KD, Coffman CJ, Golightly YM, et al (2009) Daily pain variations among patients with hand, hip, and knee osteoarthritis. Osteoarthritis Cartilage 17:1275–1282. https://doi.org/10.1016/j.joca.2009.03.021

3. Zempsky WT, Palermo TM, Corsi JM, et al (2013) Daily changes in pain, mood and physical function in children hospitalized for sickle cell disease pain. Pain Res Manag 18:33–38. https://doi.org/10.1155/2013/487060

4. Kratz AL, Murphy SL, Braley TJ (2017) Ecological Momentary Assessment of Pain, Fatigue, Depressive, and Cognitive Symptoms Reveals Significant Daily Variability in Multiple Sclerosis. Arch Phys Med Rehabil 98:2142–2150. https://doi.org/10.1016/j.apmr.2017.07.002

5. Farrar JT, Troxel AB, Haynes K, et al (2014) Effect of variability in the 7-day baseline pain diary on the assay sensitivity of neuropathic pain randomized clinical trials: an ACTTION study. Pain 155:1622–1631. https://doi.org/10.1016/j.pain.2014.05.009

6. Kuzu D, Troost JP, Carlozzi NE, et al (2022) How Do Fluctuations in Pain, Fatigue, Anxiety, Depressed Mood, and Perceived Cognitive Function Relate to Same-Day Social Participation in Individuals With Spinal Cord Injury? Arch Phys Med Rehabil 103:385–393. https://doi.org/10.1016/j.apmr.2021.07.809

7. Schneider S, Junghaenel DU, Keefe FJ, et al (2012) Individual differences in the day-to-day variability of pain, fatigue, and well-being in patients with rheumatic disease: associations with psychological variables. Pain 153:813–822. https://doi.org/10.1016/j.pain.2012.01.001

8. O'Brien EM, Waxenberg LB, Atchison JW, et al (2011) Intraindividual variability in daily sleep and pain ratings among chronic pain patients: bidirectional association and the role of negative mood. Clin J Pain 27:425–433. https://doi.org/10.1097/AJP.0b013e318208c8e4

9. Erickson BA, Herman T, Hahn AE, et al (2021) A Mobile Phone Application for Assessing Daily Variation in Pain Location and Pain Intensity in Patients with Urologic Chronic Pelvic Pain Syndrome: A MAPP Network Study. Urol Pract 8:189–195. https://doi.org/10.1097/UPJ.0000000000000203

10. Dzierzewski JM, Williams JM, Roditi D, et al (2010) Daily variations in objective nighttime sleep and subjective morning pain in older adults with insomnia: evidence of covariation over time. J Am Geriatr Soc 58:925–930. https://doi.org/10.1111/j.1532-5415.2010.02803.x

11. Wesolowicz DM, Bishop MD, Robinson ME (2021) An Examination of Day-to-Day and Intraindividual Pain Variability in Low Back Pain. Pain Med Malden Mass 22:2263–2275. https://doi.org/10.1093/pm/pnab119

12. Andersen LL, Fallentin N, Ajslev JZN, et al (2017) Association between occupational lifting and day-to-day change in low-back pain intensity based on company records and text messages. Scand J Work Environ Health 43:68–74. https://doi.org/10.5271/sjweh.3592

13. Martire LM, Wilson SJ, Small BJ, et al (2016) COMT and OPRM1 Genotype Associations with Daily Knee Pain Variability and Activity Induced Pain. Scand J Pain 10:6–12. https://doi.org/10.1016/j.sjpain.2015.07.004

14. Taylor SS, Davis MC, Zautra AJ (2013) Relationship status and quality moderate daily pain-related changes in physical disability, affect, and cognitions in women with chronic pain. Pain 154:147–153. https://doi.org/10.1016/j.pain.2012.10.004

15. Vivekanantham A, Selby D, Lunt M, et al (2023) Day-to-day variability of knee pain and the relationship with physical activity in people with knee osteoarthritis: an observational, feasibility study using consumer smartwatches. BMJ Open 13:e062801. https://doi.org/10.1136/bmjopen-2022-062801

16. Fogel A (2013) Better or Worse: a Study of Day-to-Day Changes over Five Months of Rosen Method Bodywork Treatment for Chronic Low Back Pain. Int J Ther Massage Bodyw 6:14–24. https://doi.org/10.3822/ijtmb.v6i3.200

Appendix D

Appendix to Chapter 3: Additional files from manuscript 1 submitted to *PLOS ONE*.

Additional file 1: GRIPP2 short form

Section and topic	**Item**	**Reported on page No**
1: Aim	Report the aim of PPI in the study	5
2: Methods	Provide a clear description of the methods used for PPI in the study	6–8
3: Study results	Outcomes—Report the results of PPI in the study, including both positive and negative outcomes	9–14
4: Discussion and conclusions	Outcomes—Comment on the extent to which PPI influenced the study overall. Describe positive and negative effects	14
5: Reflections/critical perspective	Comment critically on the study, reflecting on the things that went well and those that did not, so others can learn from this experience	14–15

GRIPP2 short form from: Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. BMJ. 2017 Aug 2;j3453.

Additional File 2: Charity organisations that advertised the study

Charity organisation	Method of advertisement
Action for ME	Website, Twitter, Facebook
Arthritis Action	Website
Arthritis and Musculoskeletal Alliance (ARMA)	Newsletter, Twitter
Backcare	Website
Lupus UK	Online forum
Migraine Trust	Website, Twitter, Facebook
National Axial Spondyloarthritis Society (NASS)	Website, Newsletter
National Rheumatoid Arthritis Society (NRAS)	Website, Twitter, Facebook, Instagram
People in Research	Website
Postural Tachycardia Syndrome UK (PoTS UK)	Twitter, Facebook
PsAZZ Support Group	Contacted network
Scleroderma & Raynaud's UK (SRUK)	Twitter, Facebook
The Erythromelgalgia Warriors	Website, Twitter, Facebook
UK Gout Society	Twitter
Vocal	Contacted network

Additional File 3: Planned structure of the focus group

Section of focus group	Purpose	Example questions/statements	Anticipated duration (mins)
Introduction and general overview	• Greeting • Use of Zoom • Review of ground rules • Ice breaker	• Zoom: how to use hands up function • We sent around some ground rules as suggested by the university. Is there anything that is missing that you would like to add?	20
Introduction to our ideas	• Outline the context of research • Explain proposed research	Presentation	5
Breakout rooms	• Understand initial thoughts about the research and perceived interest in different predictands	• Here are some common patterns of pain severity. Which one(s) do you relate to? • What pain features would you want to know about?	15
Break			10
Group discussion	• Bring thoughts from breakout rooms together	• What pain features did you come up with? • Were there any in other groups that you think are good that you hadn't thought of?	15
Building a questionnaire	• Outline of a questionnaire that would be meaningful	• We want to build a questionnaire to ask other people what information they would like to know about pain patterns in the future. Here are some example questions that we have thought of • Are there any questions here that you think we should remove? • Are there any questions that we should include? • What possible answers should we have? • What information would you add or remove to the images to make them clearer?	20
Conclusion	• Thank participants & let them know what the next steps are		5

Additional File 4: UK-wide questionnaire

Start of Block: Info and consent

Thank you for your interest in helping us build a pain forecast. In 2016, the Centre for Epidemiology Versus Arthritis at the University of Manchester ran a large study, collecting data about people's symptoms and comparing this to the weather. You can find out more about this study by visiting the Cloudy with a Chance of Pain website.

Using this data, we are now aiming to build a pain forecast. This will be like a weather forecast but used for predicting pain severity in the near future. Being able to forecast pain severity in this way will provide benefits for patients in the future.

During this survey we are interested in learning about the information that you would be interested in receiving about your pain severity in the near future.

The next page will outline the information that we will collect in this study, your data protection rights and how to find out more information. If you are still happy to continue, you will be asked a series of questions about our research.

Thank you for your time.

Page Break

Before you continue, it is important that you read the following information to understand what this research entails and what will happen to your information by completing this questionnaire. If you have any further questions, please use the contact details provided at the end of this page.

The research is being conducted by Claire Little (PhD student, Centre for Epidemiology Versus Arthritis, University of Manchester) and supervised by Professor John McBeth (Centre for Epidemiology Versus Arthritis, University of Manchester). It has been approved by the University of Manchester Proportionate Ethic Committee, [UREC reference number 2021-11862-19751].

To participate in this research, you must be over the age of 18, with a chronic pain condition, live in the UK, and be able to read English.

It is up to you whether you want to take part. On the next page, you consent to us using your anonymous answers for research purposes. During the questionnaire, you may stop at any time and your responses will not be used. However, once you submit the questionnaire, it will not be possible to remove your responses from the anonymised aggregated data.

Your responses will only be viewed by researchers at the University of Manchester. We will use these to understand your views on a pain forecast in our future work. It is also possible that some answers may be published or used in a student thesis. By continuing, you consent to this use.

Please note the following information in relation to the processing of your data. The University of Manchester, as Data Controller for this project takes responsibility for the protection of the personal information that this study is collecting about you. In order to comply with the legal obligations to protect your personal data the University has safeguards in place such as policies and procedures. All researchers are appropriately

trained. Data will be held securely by the research team on behalf of the University of Manchester according to the University's data protection and information security policies. Your responses to this questionnaire will be stored for 5 years, in accordance with the University's Record Retention Policy. We are collecting and storing this personal information in accordance with the General Data Protection Regulation (GDPR) and Data Protection Act 2018 which legislate to protect your personal information. The legal basis upon which we are using your personal information is "public interest task" and "for research purposes". For more information about the way we process your personal information and comply with data protection law please see our Privacy Notice for Research Participants. Complaints Procedure

If you have any complaints, you can contact Claire's supervisor, Professor John McBeth at john.mcbeth@manchester.ac.uk

If you wish to make a formal complaint to someone independent of the research team or if you are not satisfied with the response you have gained from the researchers in the first instance then please contact

The Research Governance and Integrity Officer, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester, M13 9PL,

by emailing: research.complaints@manchester.ac.uk

or by telephoning 0161 306 8089.

Contact details

If you have any further questions, please contact Claire Little at claire.little@postgrad.manchester.ac.uk.

Page Break

Have you read the information on the previous page? (Required)

- ○ Yes
- ○ No

Do you confirm that you are over 18, have a chronic pain condition, live in the UK and can read English? (Required)

- ○ Yes
- ○ No

Do you understand that participation in this study is voluntary but once you have submitted the questionnaire, it will not be possible to remove your answers from the data set? (Required)

- ○ Yes
- ○ No

Do you provide consent for aggregated responses to be used in a student thesis, reports, presentations or journals? (Required)

- ○ Yes
- ○ No

End of Block: Info and consent

Start of Block: Demographic

What is your gender? (Required)

- ○ Male
- ○ Female
- ○ Non-binary / third gender
- ○ Prefer not to say

What is your age? (Required)

- ○ 18-25
- ○ 26-35
- ○ 36-45
- ○ 46-55
- ○ 56-65

○ 65+

○ Prefer not to say

Has your doctor ever told you that you have any of the following conditions? (Required; Choose as many as relevant)

☐ Rheumatoid Arthritis

☐ Osteoarthritis

☐ Spondyloarthropathy

☐ Gout

☐ Unspecific Arthritis

☐ Fibromyalgia

☐ Chronic headache

☐ Neuropathic pain

☐ Other (please specify)

_____

Where do you generally experience pain? (Required; Choose as many as relevant)

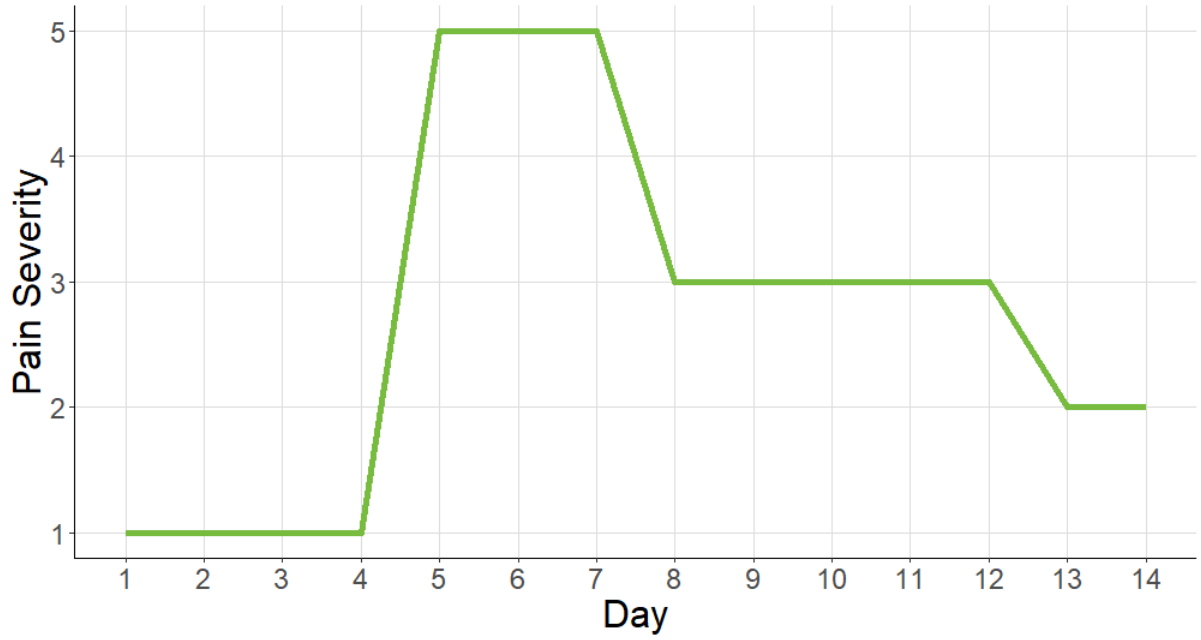
- [ ] Mouth or jaw
- [ ] Neck or shoulder
- [ ] Back
- [ ] Stomach or abdominal
- [ ] Hip
- [ ] Knee
- [ ] Hands
- [ ] Feet
- [ ] Pain at multiple sites
- [ ] Pain all over body

End of Block: Demographic

Start of Block: Block 2

We know that people's pain severity can vary every day. In 2016, we collected information about people's pain severity on a daily basis, ranging from no pain to very severe pain.

Here is an example. For the first few days, the person has reported no pain, but on day 5, they begin reporting very severe pain. Later on, the pain severity reduces to moderate and then mild.



We are hoping to use this information to produce a pain forecast, that could predict pain severity in the future. We are interested in your opinions about this pain forecast.

End of Block: Block 2

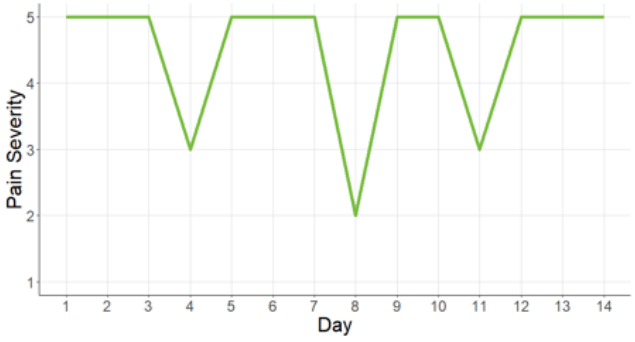
Start of Block: Block 10

Section 1: How do you experience pain?

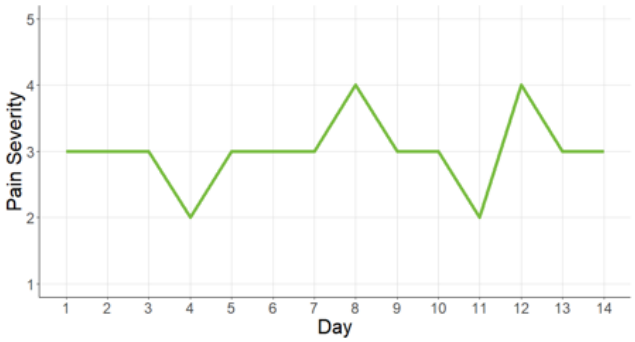
Everyone experiences pain in different ways. In this section, we are interested in understanding your experience of different pain patterns.

Some people experience pain patterns as shown by the graphs below. Other people do not experience any of these pain patterns. Which, if any, of these pain patterns do you ever experience? (Choose, by clicking on the graphs, as many as required)

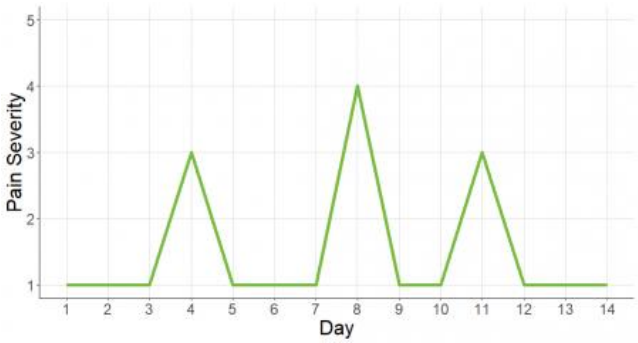
☐ Occasional decreases in pain severity that last a very short time (e.g. one day)



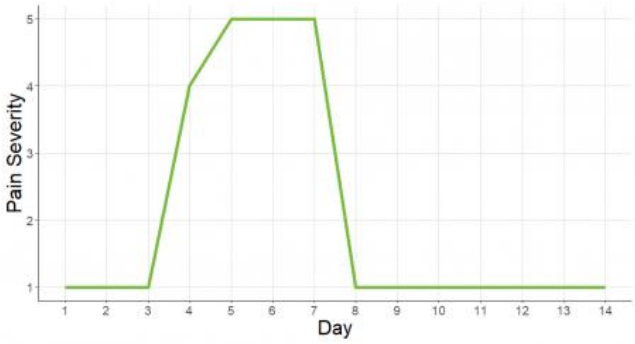
☐ Variable pain severity that changes day-to-day



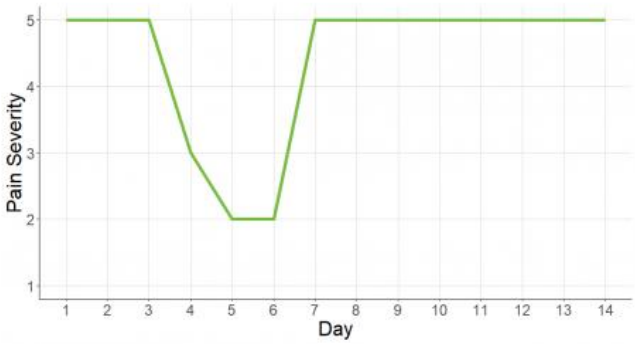
☐ Occasional increases in pain severity that last a very short time (e.g. one day)



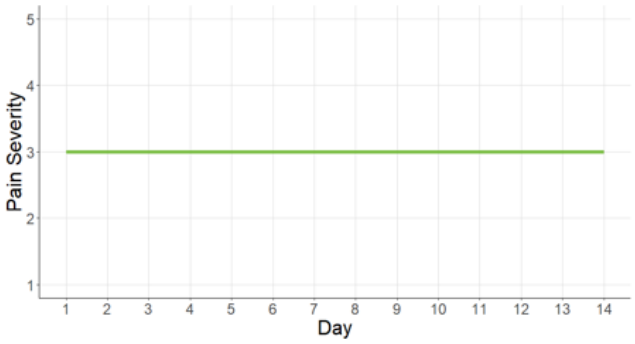
☐ Rapid increase in pain severity that lasts a few days (pain flare)



- [ ] Generally high pain with a period of lower pain



- [ ] Constant pain that doesn't tend to change



- [ ] None of these represent pain that I experience
- [ ] I experience a different pain pattern. (Please describe below).

_____

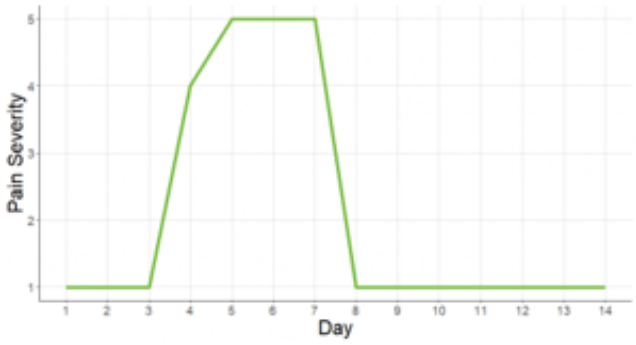
Page Break

Section 2: What information do you want from a pain forecast?

We are using pain patterns (like the ones on the previous page) to build a **pain forecast**, that would predict your pain severity in the future. We are interested in understanding the information that you would like from a pain forecast.

To start with, please think very generally about the type of information that you would like from a pain forecast. Which of the following would you like a pain forecast to provide for you? (Choose as many relevant)

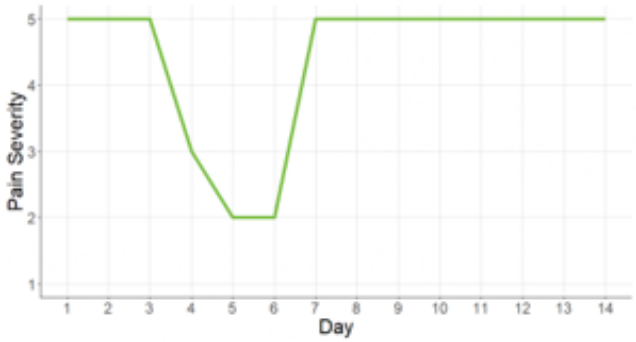
- [ ] Information about a pain flare (e.g. timing/duration/severity)



- [ ] Information about the fluctuations in pain severity (e.g. how stable pain severity will be over the week)

- [ ] Information about pain severity on a scale of 1 (no pain) to 5 (very severe pain) (e.g. what pain severity will be on this scale tomorrow)

☐ Information about a period of low/no pain severity (e.g. timing/duration)

- 

☐ Other: _____

If we could predict a period of low pain severity, what specific information would you want to know? (Please rank the responses from most wanted information to least wanted information by dragging the responses to be in the correct order)

- _____ When might the period of low pain severity start?
- _____ When might the period of low pain severity end?
- _____ How many days might the period of low pain severity last for?
- _____ Might my pain-related quality of life also be better?
- _____ Might my other symptoms also be better? (e.g. fatigue, morning stiffness, sleep quality)

Is there any other information about periods of low pain severity that you would like to know about? (Optional)

_____

If we could predict a pain flare, what specific information would you want to know? (Please rank the responses from most wanted information to least wanted information, by dragging the responses to be in the correct order)

- _____ When might the pain flare start?
- _____ When might the pain flare end?
- _____ How many days might the pain flare last for?
- _____ Might my pain-related quality of life be lower?
- _____ Might my other symptoms vary? (e.g. fatigue, morning stiffness, sleep quality)
- _____ How severe will the pain flare be?

Is there any other information about pain flares that you would like to know about? (Optional)

_____

End of Block: Block 10

Start of Block: Block 4

Section 3: How would you use a pain forecast?

Now that you have considered the type of information that a pain forecast could provide, we are interested in how you would use a pain forecast.

On the previous page, we asked you to consider information that you might want from a pain forecast. If a pain forecast could provide useful information for you, do you think that you would use a pain forecast?

- ○ Definitely yes
- ○ Probably yes
- ○ Might or might not
- ○ Probably not
- ○ Definitely not

If so, what would you use a pain forecast for? (Choose as many as required)

- ☐ To know when my pain severity might be better/worse
- ☐ To understand the triggers of my pain
- ☐ To help choose which medication to take
- ☐ To help plan non-pharmacological interventions (i.e. other methods of helping pain that aren't medication)
- ☐ To help make plans (e.g. shopping, social)
- ☐ Other _____

End of Block: Block 4

Additional File 5 – Sensitivity analysis

In this appendix, we acknowledge that a large proportion of our respondents reported fibromyalgia (46%) and osteoarthritis (33%) as a chronic pain condition. We conducted sensitivity analyses to compare (1) the responses between those respondents who reported fibromyalgia as a pain condition and those that did not, and (2) the responses between those respondents who reported osteoarthritis as a pain condition and those that did not.

For each question, we report the number and percentage of respondents in the disease subgroups who selected each response and perform a chi-squared test to test whether these responses are significantly different from these not in the corresponding subgroup.

Table S1: Responses to the question: "Which of the following would you like a pain forecast to provide for you?" Participants could select more than one option.

Pre-specified response	**Number (and percentage) of full population who selected response**	**Number (and percentage) of subgroup with fibromyalgia who selected response**	**Number (and percentage) of subgroup with osteoarthritis who selected response**
Information about a pain flare	100 (67.6%)	47 (69.1%)	34 (69.4%)
Information about fluctuations in pain severity	94 (63.5%)	47 (69.1%)	27 (55.1%)
Information about pain severity on a scale of 1 to 5	70 (47.3%)	30 (44.1%)	23 (46.9%)
Information about a period of low/no pain severity	51 (34.5%)	25 (36.8%)	15 (30.6%)
Other/None	13 (8.8%)	5 (7.4%)	5 (10.2%)

Results to the question: "Which of the following would you like a pain forecast to provide for you?" are reported in Table S1. A chi squared test of the responses for this question from participants who did not report fibromyalgia against those who did report fibromyalgia gave a p-value of 0.2414. For the same question, a chi squared test of the responses from participants who did not report osteoarthritis against those who did report osteoarthritis gave a p-value of 0.2202. Therefore, there is no evidence that the subgroups gave statistically significantly different responses to the population.

Table S2: Responses to the question: "If a pain forecast could provide useful information for you, do you think that you would use a pain forecast?"

Response	**Number (and percentage) of full population who selected response**	**Number (and percentage) of subgroup with fibromyalgia who selected response**	**Number (and percentage) of subgroup with osteoarthritis who selected response**
Definitely not	0 (0%)	0 (0%)	0 (0%)
Probably not	11 (7.4%)	2 (2.9%)	4 (8.2%)
Might or might not	24 (16.2%)	9 (13.2%)	8 (16.3%)
Probably yes	63 (42.6%)	27 (39.7%)	16 (32.7%)
Definitely yes	50 (33.8%)	30 (44.1%)	21 (42.9%)

Results to the question: "If a pain forecast could provide useful information for you, do you think that you would use a pain forecast?" are reported in Table S2. A chi squared test of the responses from participants who did not report fibromyalgia against those

who did report fibromyalgia gave a p-value of 0.2133. For the same question, a chi squared test of the responses from participants who did not report osteoarthritis against those who did report osteoarthritis gave a p-value of 0.2133. Therefore, there is no evidence that the subgroups gave statistically significantly different responses to the population.

Table S3: Responses to the question "What would you use a pain forecast for?" Participants could select more than one option.

Pre-specified response	Number (and percentage) of full population who selected response	Number (and percentage) of subgroup with fibromyalgia who selected response	Number (and percentage) of subgroup with osteoarthritis who selected response
To help make plans (e.g. shopping, social)	123 (83.1%)	59 (86.8%)	45 (91.8%)
To understand the triggers of my pain	113 (76.4%)	53 (77.9%)	38 (77.6%)
To know when my pain severity might be better/worse	92 (62.2%)	46 (67.6%)	28 (57.1%)
To help plan non-pharmacological interventions	70 (47.3%)	29 (42.6%)	22 (44.9%)
To help choose which medication to take	46 (31.1%)	24 (35.3%)	21 (42.9%)
Other/None	16 (10.8%)	6 (8.8%)	5 (10.2%)

The responses to the question: "What would you use a pain forecast for?" are reported in Table S3. A chi squared test of the responses from participants who did not report fibromyalgia against those who did report fibromyalgia gave a p-value of 0.2243. For the same question, a chi squared test of the responses from participants who did not report osteoarthritis against those who did report osteoarthritis gave a p-value of 0.2243. Therefore, there is no evidence that the subgroups gave statistically significantly different responses to the population.

Appendix E

Appendix to Chapter 4: Supplementary materials from manuscript 2 submitted to *JMIR.*

Supplementary Materials 1 (Co-morbid pain conditions)

Table S1: Reported co-morbid pain conditions by condition

Condition	Participants reporting condition: *n*	Participants reporting co-morbid pain condition: *n (%)*
Rheumatoid arthritis	548	299 (54.6)
Osteoarthritis	975	720 (73.8)
Spondyloarthropathy	254	196 (77.2)
Gout	96	91 (94.8)
Unspecific arthritis	1028	688 (66.9)
Fibromyalgia	718	577 (80.4)
Chronic headache	274	237 (86.5)
Neuropathic pain	427	379 (88.8)
Other/no medical diagnosis	667	381 (57.1)

Table S2: Most commonly reported co-morbid conditions

Conditions	Participants reporting co-morbid pain conditions: *n*
Osteoarthritis and Unspecified Arthritis	386
Osteoarthritis and Fibromyalgia	210
Unspecified Arthritis and Fibromyalgia	206
Fibromyalgia and Neuropathic Pain	193
Fibromyalgia and Other Pain Condition	187
Osteoarthritis and Neuropathic Pain	154
Rheumatoid Arthritis and Osteoarthritis	150
Unspecified Arthritis and Neuropathic Pain	144
Osteoarthritis and Other Pain Condition	133
Unspecified Arthritis and Other Pain Condition	123

The ten most common combinations of comorbid conditions are presented with the number of people reporting both conditions

Supplementary Materials 2 (Sensitivity Analyses 1 & 2)

This section reports the results from two of the sensitivity analyses. In the main analysis, trajectories were compared using the Manhattan distance and clustered using the *k*-medoids algorithm. In the first sensitivity analysis, trajectories were compared using the Euclidean distance and clustered using the KmL package in R. In the second sensitivity analysis, trajectories were summarised in a one-hot encoded feature vector, compared using the Jaccard distance and clustered using the k-medoids algorithm.

For each of the sensitivity analyses, the methods of the main analysis are repeated. The results of these analyses are presented alongside the results for the main analysis.

Identifying optimal number of clusters

First, remaining variability within clusters (within-cluster sum of squares) was calculated for *k* (number of clusters) between 1 and 20. The results are presented in figure S1. There is an elbow at $k = 4$ for the first sensitivity analysis, with a four-cluster solution describing 60.0% of the observed variability. There is an elbow at $k = 5$ for the second sensitivity analysis, with a five-cluster solution describing 57.0% of the variability. In the following sections, both a four-cluster solution and five-cluster solution are explored for the second sensitivity analysis.

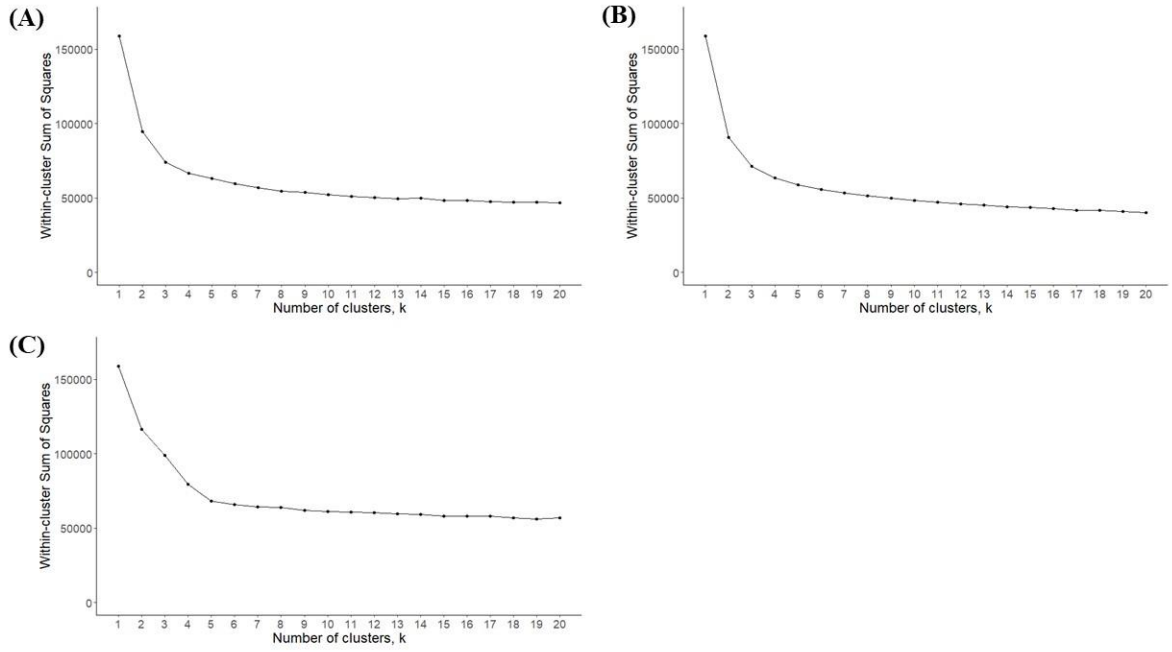
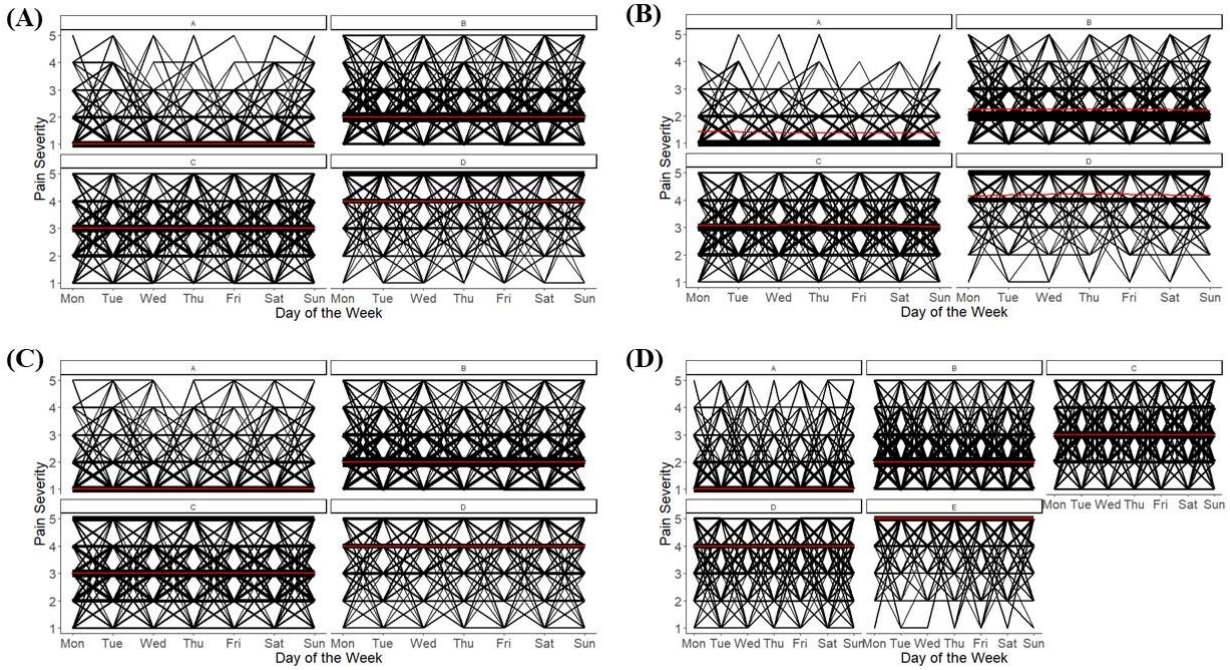


Figure S1: Unexplained variability among different cluster (k) solutions for (A) the main analysis, (B) the first sensitivity analysis and (C) the second sensitivity analysis.

Description of clusters

Four solutions are described in detail. These are: the four-cluster solution from the main analysis, the four-cluster solution from the first sensitivity analysis, the four-cluster solution from the second sensitivity analysis and the five-clusters solution from the second sensitivity analysis. For each solution, trajectories were assigned to a cluster and weighted spaghetti plots of trajectories were created to visualise the cluster solutions (Figure S1). The median or mean trajectory is provided for solutions of a k-medoids algorithm or KmL algorithm respectively. All four-cluster solutions could be described with the clusters A = “low or no pain”, B = “mild pain”, C = “moderate pain”, D = “severe pain”. The five-cluster solution could be described with the same descriptors and E = “very severe pain”.



Cluster C (moderate)	36.2	36.3	40.9	36.6
Cluster D (severe)	16.3	13.4	13.3	11.0
Cluster E (very severe)	NA	NA	NA	4.5

Figure S2: Weighted spaghetti plots of trajectories assigned to each cluster in (A) main analysis, (B) first sensitivity analysis, (C) four-cluster solution of second sensitivity analysis, (D) five-cluster solution of second sensitivity analysis. All cluster solutions can be described by A = "low or no pain", B = "mild pain", C = "moderate pain", D = "severe pain", with (D) also containing E = "very severe pain".

Table S3: Percentage of trajectories assigned to each cluster in most optimal cluster solutions

The percentage of trajectories assigned to each cluster are shown in Table S3. Compared to the main analysis, the first sensitivity analysis had a larger percentage of trajectories in Cluster A, and a lower percentage in Cluster D. Compared to the main analysis, the second sensitivity analysis had a smaller percentage of trajectories in Cluster D, in both the four- and five-cluster solutions. In the five-cluster solution, 4.5% of trajectories were assigned to Cluster E.

Comparison of trajectory assignment

Each trajectory is assigned to a cluster in the four-cluster solutions of the main analysis, first sensitivity analysis and second sensitivity analysis. For the trajectories assigned to each cluster in the main analysis, Table S2 compares the percentage of trajectories assigned to the same cluster in the first sensitivity analysis and the second sensitivity analysis. Of the trajectories assigned to Cluster A in the main analysis, at least 97% of trajectories are also assigned to Cluster A in both sensitivity analyses. However, of the trajectories assigned to Cluster D in the main analysis, only 80.3% and 77.4% are still assigned to Cluster D in the first sensitivity analysis and second sensitivity analysis respectively. This suggests that a large proportion of trajectories are assigned to the same cluster in each of the analyses, but a significant proportion may be assigned to different clusters, perhaps reflecting the changes in cluster sizes among the different analyses.

Table S2: Comparison of trajectories assigned to the same cluster as the main analysis

Cluster in main analysis	Number (percentage) same in first sensitivity analysis	Number (percentage) same in second sensitivity analysis
Cluster A (low/no pain)	1672 (97.5%)	1714 (100%)
Cluster B (mild)	7238 (87.8%)	7822 (94.9%)
Cluster C (moderate)	7108 (84.9%)	7886 (94.1%)
Cluster D (severe)	2877 (80.3%)	2775 (77.4%)

Supplementary Materials 3 (Sensitivity Analysis 3)

This section reports the results from the final sensitivity analysis. The main analysis defined complete participant weeks for each individual as those containing complete pain severity data for each day in a Monday-Sunday week. Six alternative definitions are explored here, using complete pain severity data for each day in (1) Tuesday-Monday weeks, (2) Wednesday-Tuesday weeks, (3) Thursday-Wednesday weeks, (4) Friday-Thursday weeks, (5) Saturday-Friday weeks and (6) Sunday-Saturday weeks.

For each definition, all complete participant weeks were identified, compared using the Manhattan distance and clustered using the CLARA program of the *k*-medoids algorithm. The remaining variability within clusters (within-cluster sum of squares) was calculated for *k* (number of clusters) between 1 and 20. The results are presented in Figure S3, for each data set separately. A four-cluster solution is optimal for each dataset. The percentage of trajectories assigned to each cluster are reported in Table S4. The graphs in Figure S3 and cluster sizes in Table S4 show similarity in the results for trajectories starting on different days of the week, suggesting that the main analysis is robust to the day of the week on which the trajectory began.

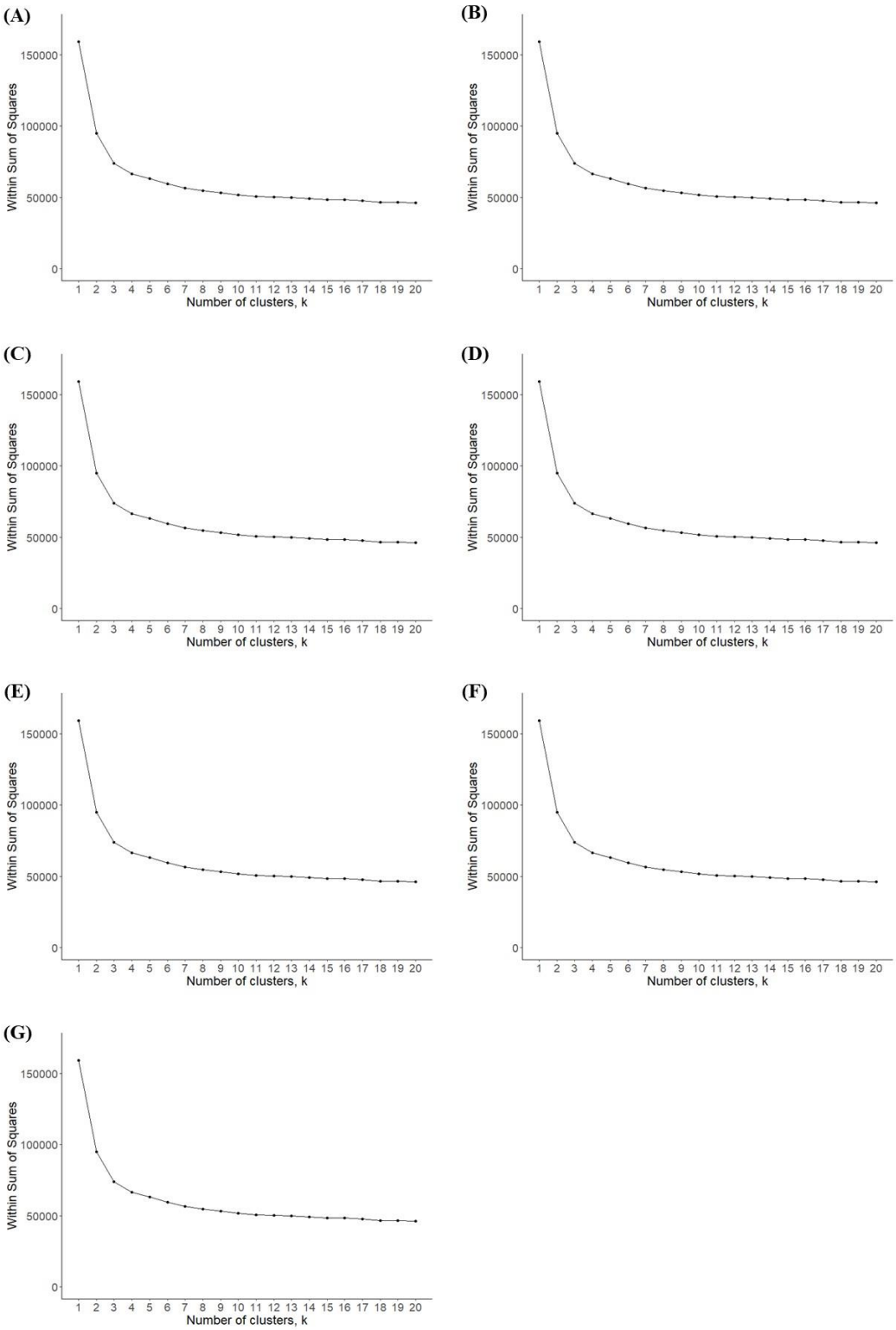


Figure S3: Unexplained variability among different cluster (k) solutions using trajectories with complete data (A) Monday – Sunday, (B) Tuesday – Monday, (C) Wednesday – Tuesday, (D) Thursday – Wednesday, (E) Friday – Thursday, (F) Saturday – Friday, (G) Sunday – Saturday.

Table S4: Comparison of cluster size among four-cluster solution for trajectories starting on different days of the week

First day of trajectories	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Number of trajectories in analysis	21919	22103	22320	22653	22404	22255	22067
Percentage in Cluster A (low/no pain)	7.8%	7.8%	7.9%	7.8%	7.9%	7.7%	7.7%
Percentage in Cluster B (mild pain)	37.6%	37.7%	37.7%	37.6%	37.6 %	37.7%	37.5%
Percentage in Cluster C (moderate pain)	38.2%	38.2%	38.3%	38.2%	38.0 %	38.2%	38.4%
Percentage in Cluster D (severe pain)	16.3%	16.3%	16.2%	16.4%	16.4 %	16.4%	16.3%

Supplementary Materials 4 (Demographic data in transition analysis)

Table S5: Comparison of demographics among participants included in the main analysis and participants included in the transition analysis.

		Number (percentage) of participants in main analysis	Number (percentage) of participants in transition analysis
Age	17–24	67 (2.4)	30 (1.7)
	25–34	255 (9.1)	139 (7.9)
	35–44	508 (18.1)	296 (16.8)
	45–54	755 (26.9)	468 (26.6)
	55–64	788 (28.1)	528 (30.0)
	65–86	434 (15.5)	300 (17.0)
Sex	Female	2333 (83.1)	1467 (83.3)
	Male	474 (16.9)	294 (16.7)
Condition*	Rheumatoid arthritis	548 (19.5)	365 (20.7)
	Osteoarthritis	975 (34.7)	644 (36.6)
	Spondyloarthropathy	254 (9.0)	161 (9.1)
	Gout	96 (3.4)	60 (3.4)
	Unspecific arthritis	1028 (36.6)	676 (38.4)
	Fibromyalgia	718 (25.6)	446 (25.3)
	Chronic headache	274 (9.8)	172 (9.8)
	Neuropathic pain	427 (15.2)	253 (14.4)

	Other/no medical diagnosis	668 (23.8)	402 (22.7)

[a]Percentages exceed 100% because participants could report multiple chronic-pain conditions.

Appendix F

Appendix to Chapter 5: Supplementary materials from manuscript 3 submitted to *Lancet Digital Health.*

Table S1: Demographic data of weeks included as complete weeks and those that were available but not complete.				
		Data of complete weeks	Data of available weeks that were not complete	*p*-value of differences between groups
Sex (%)				
	Female	82.7	82.1	0.237
	Male	17.3	17.9	0.237
Age (mean)		53.0	53.6	**0.002**
Chronic pain condition (%)				
	Rheumatoid arthritis	19.2	19.6	0.557
	Osteoarthritis	41.0	40.0	0.151
	Spondyloarthropathy	8.6	9.0	0.243
	Gout	3.7	3.0	**0.005**
	Unspecific arthritis	43.6	40.2	**<0.001**
	Fibromyalgia	22.8	24.2	**0.019**
	Chronic headache	7.3	7.9	0.095
	Neuropathic pain	14.4	15.0	0.246
	Other/no medical diagnosis	21.0	21.3	0.587
Site of pain (%)				
	Head	13.6	15.3	**<0.001**
	Face	4.4	5.8	**<0.001**
	Mouth/jaws	12.1	14.1	**<0.001**
	Neck/shoulder	57.9	59.4	**0.030**
	Back	55.9	55.0	0.239
	Stomach/abdominal	13.0	11.9	**0.023**
	Hip	51.7	49.1	**<0.001**
	Knee	66.4	63.1	**<0.001**
	Hands	63.8	62.5	**0.046**
	Feet	47.9	47.1	0.264
	Multisite	43.2	42.0	0.084
	All	14.6	13.8	0.111

Table S2: Number (%) of available weeks in which listed variables contained missing data	
Fatigue	4168 (19.0%)
Mood	4158 (19.0%)
Morning stiffness	4640 (21.2%)
Pain impact	4327 (19.7%)
Patient wellbeing	4276 (19.5%)
Exercise	4554 (20.8%)
Sleep quality	4856 (22.2%)
Time spent outside	4679 (21.3%)
Waking up tired	4787 (21.8%)
Daily mean windspeed	1553 (7.1%)
Daily mean temperature	995 (4.5%)
Daily mean dewpoint	995 (4.5%)
Daily mean pressure	995 (4.5%)
Daily mean relative humidity	995 (4.5%)

Available weeks were incomplete if they contained missing data for at least one of the listed variables. The number (%) of available weeks with missing data is reported. Available weeks may have contained missing data for multiple variables, and therefore weeks may be counted multiple times.

Table S3: Probability of pain recovery and pain worsening, at baseline, and given one-unit changes in significant variables, by cluster.

	Cluster A		Cluster B		Cluster C		Cluster D	
	P(Pain recovery)	P(Pain worsening)	P(Pain recovery)	P(Pain worsening)	P(Pain recovery)	P(Pain worsening)	P(Pain recovery)	P(Pain worsening)
Baseline	0.128	0.120	0.172	0.163	0.202	0.192	0.182	0.180
Worsened fatigue by 1-unit	0.117	0.131	0.151	0.186	0.173	0.223	0.135	0.235
Improved fatigue by 1-unit	0.139	0.109	0.196	0.142	0.235	0.163	0.237	0.134
Worsened mood by 1-unit	0.122	0.126	0.164	0.172	0.187	0.207	0.167	0.195
Improved mood by 1-unit	0.134	0.114	0.181	0.155	0.218	0.177	0.197	0.166
More exercise by 1-unit	.	.	0.165	0.171	0.193	0.201	0.172	0.191
Less exercise by 1-unit	.	.	0.180	0.156	0.212	0.183	0.192	0.171
Reduced time spent outside by 1-unit	.	.	.	.	0.211	0.183	0.192	0.171
Increased time spent outside by 1-unit	.	.	.	.	0.193	0.200	0.172	0.190
Worsened stiffness by 1-unit	0.063	0.217	0.112	0.239	0.139	0.267	0.128	0.246
Improved stiffness by 1-unit	0.229	0.058	0.250	0.105	0.280	0.131	0.248	0.127
Worsened lag stiffness by 1-unit	.	.	0.167	0.169	.	.	.	.
Improved lag stiffness by 1-unit	.	.	0.178	0.158	.	.	.	.
Worsened sleep by 1-unit	0.122	0.126	0.165	0.170	0.196	0.198	0.179	0.183
Improved sleep by 1-unit	0.134	0.114	0.180	0.156	0.209	0.185	0.185	0.177

Worsened pain inteference by 1-unit	0.020	0.395	0.065	0.340	0.097	0.343	0.097	0.301
Improved pain inteference by 1-unit	0.411	0.018	0.354	0.061	0.357	0.090	0.303	0.096
Worsened lag pain inteference by 1-unit	.	.	0.167	0.169	0.192	0.202	.	.
Improved lag pain inteference by 1-unit	.	.	0.178	0.158	0.213	0.181	.	.
Worsened wellbeing by 1-unit	0.110	0.139	0.142	0.196	0.165	0.232	0.149	0.218
Improved wellbeing by 1-unit	0.148	0.102	0.207	0.134	0.244	0.156	0.219	0.148
Increased lag windspeed by 1-unit	.	.	.	.	0.208	0.186	.	.
Decreased lag windspeed by 1-unit	.	.	.	.	0.197	0.197	.	.

Appendix G

Appendix to Chapter 6: Supplementary materials from manuscript 4 prepared for submission to *npj Digital Medicine.*

Table S1: Hyperparameters λ and α for the multinomial models fit for each cluster

		Cluster A	Cluster B	Cluster C	Cluster D
λ	Model 1	3.0×10^{1}	4.4×10^{1}	6.6×10^{-3}	4.5×10^{-3}
	Model 2	9.6×10^{-4}	8.4×10^{1}	5.0×10^{-2}	5.1×10^{-3}
	Model 3	2.2×10^{-2}	8.4×10^{1}	1.1×10^{-1}	2.2×10^{-4}
	Model 4	**2.7×10^{-1}**	8.4×10^{1}	**2.5×10^{-2}**	**3.1×10^{-5}**
	Model 5	2.7×10^{-2}	**3.5×10^{-3}**	5.8×10^{-2}	1.0×10^{-1}
α	Model 1	0	0	0	0.7
	Model 2	0.3	0	0	0.2
	Model 3	0.8	0	0	0.8
	Model 4	**0**	0	**0.1**	**0.8**
	Model 5	0.8	**0.7**	0.3	0

Hyperparameters λ and α are reported for each multinominal model: α = 0 represents pure ridge regression, α = 1 represents pure LASSO, and 0 < α < 1 represents a combination of ridge regression and LASSO. λ impacts the size of the overall elastic net penalty term, with larger values of λ indicating a larger penalty term, and hence more shrinkage of the coefficients towards, or to, zero.

Hyperparameters of the optimal model for each cluster and marked in bold text.

New candidate predictors in Model 1 = Sex, number of pain conditions, number of pain sites, age.

New candidate predictors in Model 2 = Pain severity in origin week: maximum, minimum, PAC, MSSD.

New candidate predictors in Model 3 = Wind-speed, temperature, dewpoint temperature, pressure, relative humidity in origin week: mean, SD.

New candidate predictors in Model 4 = Fatigue, mood, morning stiffness, pain interference, well-being, physical activity, sleep quality, time spent outside, waking up tired in origin week: median, PAC, MSSD.

New candidate predictors in Model 5 = Wind-speed, temperature, dewpoint temperature, pressure, relative humidity in destination week: mean, SD.

Table S2: Coefficients of optimal models														
	Origin cluster													
	A			B				C				D		
	Destination cluster			Destination cluster				Destination cluster				Destination cluster		
	A	B	C/D	A	B	C	D	A	B	C	D	A/B	C	D
Intercept	0.680	0.922	-1.602	-2.368	1.867	1.207	-0.706	-2.137	0.851	1.450	-0.164	-0.215	0.623	-0.408
Baseline covariates														
Sex	-0.090	0.047	0.043	0.153	-0.106	0.010	-0.058	0	0.018	-0.018	0	-0.807	0.068	0
Number of pain conditions	0.044	-0.098	0.054	-0.003	0	0.191	0	-0.035	-0.067	0.102	0	-0.981	0.010	0
Number of pain sites	-0.060	0.017	0.044	-0.036	0.032	0	0	-0.109	-0.029	0.027	0.111	-0.069	0	0.067
Age	-0.001	0	0	0	0.003	0.001	-0.013	0	0.004	0	-0.011	0.003	0	-0.010
Summary measures of pain severity from previous week														
Maximum	-0.092	0.062	0.030	-0.202	-0.370	0.149	0.423	-0.091	-0.328	0	0.489	0	-0.034	0.435
Minimum	NA	NA	NA	-1.864	0	0.350	0.029	-0.461	-0.491	0.270	0.682	-1.148	0	1.168
PAC	-0.075	0.052	0.024	0	-0.077	0.189	0	0	0.053	-0.068	0	0.106	0	-0.106
MSSD	-0.039	0.009	0.030	0	0	-0.188	0	0.050	-0.038	0	0	-0.527	0	0.311

Summary measures of weather from previous week														
Wind-speed: mean	0	-0.041	0.041	0	0.028	0	0	0	0	0.013	-0.022	-0.122	0.112	0
Wind-speed: SD	0.012	-0.051	0.039	0	0	0.082	0	0	0	0.005	0	0	-0.022	0.073
Temperature: mean	0.038	-0.030	-0.009	0	0	0.004	0	0	0.022	-0.033	0	3.539	0	-1.978
Temperature: SD	0.010	0.009	-0.018	0	0	0	0.012	0	-0.021	0.012	0	0.188	-0.146	0
Dewpoint temperature: mean	0.038	-0.034	-0.004	0	0	0	0	0	0.010	-0.006	0	-3.061	0	1.905
Dewpoint temperature: SD	0.055	-0.047	-0.008	0.055	-0.024	-0.042	0.011	0	0.063	0	-0.032	-0.156	0.235	0
Pressure: mean	0.019	-0.026	0.006	0.015	0	0	0	0	-0.001	0	0.043	-0.220	0	0.225
Pressure: SD	-0.026	0.021	0.005	0	0	-0.032	0	0.002	0	-0.012	0	-0.016	0.010	0
Relative humidity: mean	-0.001	-0.014	0.015	0.045	0	-0.101	0	0	0	0.027	0	1.376	0	-0.456
Relative humidity: SD	-0.019	0.026	-0.007	0	0	-0.018	0.067	0.028	0	-0.003	0	0.044	-0.229	0
Summary measures of self-reported variables from previous week														
Fatigue: median	-0.142	0.044	0.098	-0.407	-0.069	0.035	0.441	-0.130	-0.069	0.035	0.164	0	-0.073	0.184
Fatigue: PAC	-0.059	0.012	0.046	-0.010	0	0.018	0	0.004	0	0	0	0.056	0	-0.090

Fatigue: MSSD	0.007	-0.061	0.054	0	0	0	0	0.073	0	-0.056	0	-0.001	0	0.053
Mood: median	0.018	-0.034	0.016	-0.066	0	0.029	0	0	-0.022	0	0	0.220	-0.225	0
Mood: PAC	-0.006	-0.018	0.024	0.012	0	0	0	0	0	0.022	-0.012	0.344	0	-0.271
Mood: MSSD	-0.006	-0.022	0.028	0	-0.021	0	0	0	0	-0.020	0	-0.185	0	0.290
Morning stiffness: median	-0.184	0.180	0.003	-0.181	0	0.024	0.076	-0.022	-0.197	0.018	0.202	0.748	-0.026	0
Morning stiffness: PAC	-0.054	0.021	0.034	0.041	-0.050	0	0	0	0.046	-0.003	0	0.336	0	-0.094
Morning stiffness: MSSD	-0.033	-0.004	0.038	0	-0.016	0.016	0	0	0	0	0.069	0.006	-0.062	0
Pain interference: median	-0.037	-0.174	0.210	-0.427	0	0.240	0	-0.062	-0.170	0.115	0.117	-0.003	0	0.078
Pain interference: PAC	-0.054	0.039	0.015	-0.038	0	0	0.036	0	0	0	-0.047	-0.375	0.131	0
Pain interference: MSSD	-0.001	-0.013	0.014	0	0	-0.008	0.161	0.003	-0.002	-0.021	0.019	0.459	-0.031	0
Wellbeing: median	-0.006	0.010	-0.005	0	0	0	0	0.003	0	0	-0.059	-0.376	0.046	0
Wellbeing: PAC	-0.017	-0.017	0.034	0	-0.040	0.041	0	0	0	0.005	-0.034	0	-0.018	0.107
Wellbeing: MSSD	0.003	-0.013	0.010	0	0.036	-0.008	0	0.029	0	-0.054	0	-0.046	0.183	0
Physical activity: median	-0.028	0.014	0.014	0.034	-0.029	0.029	-0.034	0.036	-0.016	-0.037	0.017	-0.234	0.029	0
Physical activity: PAC	-0.012	0.024	-0.012	0	0	0	0	0	0	-0.005	0	-0.310	0	0.092
Physical activity: MSSD	0.002	0.011	-0.013	0.045	0	0	0	0	0	0	0	0.435	0	0
Sleep quality: median	0.076	-0.032	-0.044	0.212	0	0	0	0.058	0	-0.024	0	0.315	0	-0.149
Sleep quality: PAC	-0.007	-0.006	0.014	0	-0.017	0	0	0	0	0	-0.003	-0.183	0.041	0
Sleep quality: MSSD	-0.016	0.009	0.007	0	0	0.053	0	0.061	-0.023	-0.027	0	0.136	0	-0.086
Time spent outside: median	-0.029	0.008	0.021	0	-0.014	0.067	0	0.052	-0.035	0	0	0.135	0	-0.005
Time spent outside: PAC	0.030	-0.025	-0.005	-0.057	0.008	0	0	0	0	-0.006	0.064	-0.230	0	0.008

Time spent outside: MSSD	-0.019	0.015	0.005	0	0	0	0	0	0	0	0	-0.384	0.060	0
Waking up tired: median	-0.101	0.061	0.040	0	-0.076	0.041	0	-0.005	-0.055	0.019	0.041	0.199	-0.028	0
Waking up tired: PAC	0	-0.003	0.003	0	0	0	0	0	0	-0.002	0	0.173	0	-0.277
Waking up tired: MSSD	0.030	-0.048	0.017	-0.059	0.043	0	0	0	0	0	0	0.094	-0.224	0
Summary measures of weather in predicted week														
Wind-speed: mean	NA	NA	NA	0	0	0.070	0	NA	NA	NA	NA	NA	NA	NA
Wind-speed: SD	NA	NA	NA	0.029	0	0	0	NA	NA	NA	NA	NA	NA	NA
Temperature: mean	NA	NA	NA	0	0.047	0	-0.085	NA	NA	NA	NA	NA	NA	NA
Temperature: SD	NA	NA	NA	0	0.001	0	0	NA	NA	NA	NA	NA	NA	NA
Dewpoint temperature: mean	NA	NA	NA	0.048	0	0	0	NA	NA	NA	NA	NA	NA	NA
Dewpoint temperature: SD	NA	NA	NA	0	-0.019	0.043	0	NA	NA	NA	NA	NA	NA	NA
Pressure: mean	NA	NA	NA	0.027	0	-0.021	0	NA	NA	NA	NA	NA	NA	NA
Pressure: SD	NA	NA	NA	0	0	0	0.031	NA	NA	NA	NA	NA	NA	NA
Relative humidity: mean	NA	NA	NA	0	-0.064	0	0.108	NA	NA	NA	NA	NA	NA	NA
Relative humidity: SD	NA	NA	NA	0	0.049	0	-0.032	NA	NA	NA	NA	NA	NA	NA