Understanding the Longitudinal Characteristics of Chronic Pain in Arthritis: The Role of Intensive Longitudinal Methods in Analysing Pain and the Attributable Burden of Persistent Pain on Treatment and Health Outcomes

Dr Huai Leng (Jessica) Pisaniello

MBBS, FRACP

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Medicine

Adelaide Medical School,

Faculty of Health and Medical Sciences,

University of Adelaide

December 2023

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

The author acknowledges that copyright of published works contained within the thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Dr Huai Leng (Jessica) Pisaniello (candidate)
December 2023

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

ACPAs	Anti-Cyclic Citrullinated Peptide Antibodies
ACR	American College of Rheumatology
AI	Artificial Intelligence
AIMS	Arthritis Impact Measurement Scales
ANA	Anti-Nuclear Antibody
app	Application
AQoL	Assessment of Quality of Life
ARA	Australian Rheumatology Association
ARAD	Australian Rheumatology Association Database
BPI	Brief Pain Inventory
b/tsDMARD	Biologic/Targeted Synthetic Disease Modifying Anti-Rheumatic Drugs
CDAI	Clinical Disease Activity Index
CI	Confidence Interval
CNS	Central Nervous System
COD	Causes of Death
CPGQ	Chronic Pain Grade Questionnaire
CRP	C-Reactive Protein
csDMARD	Conventional Synthetic Disease Modifying Anti-Rheumatic Drugs
CWP	Chronic Widespread Pain
DAS28	28-joint Disease Activity Score
DAS28-P	28-joint Disease Activity Score-Proportion of Patient-Reported Components (Joint Tenderness and Visual Analogue Score)
DASH	Disabilities of the Arm, Shoulder, and Hand Questionnaire
DMARDs	Disease Modifying Anti-Rheumatic Drugs
EMA	Ecological Momentary Assessment
EQ-5D	European Quality of Life-5 Dimensions
ESM	Experience Sampling Method

ESR	Erythrocyte Sedimentation Rate
EULAR	European Alliance of Associations for Rheumatology (formerly known as the European League Against Rheumatism)
EuroQoL	European Quality of Life
FM	Fibromyalgia
GPS	Geo-Positioning System
HAQ-DI	Health Assessment Questionnaire Disability Index
HIV	Human Immunodeficiency Virus
HMM	Hidden Markov Model
HOOS	Hip Disability and Osteoarthritis Outcome Score
IASP	International Association of Pain
ICD-10	International Classification of Diseases 10th Revision
ICOAP	Measure of Intermittent and Constant Osteoarthritis Pain
IL-1	Interleukin-1
IL-6	Interleukin-6
iOS	iPhone Operating System
IQR	Interquartile Range
IRAS	Integrated Research Application System
ISD	Integrated Surface Database
IT	Information Technology
IV	Intravenous
KOALAP	Knee Osteoarthritis: Linking Activity and Pain
LBP	Low Back Pain
mHealth	Mobile Health
mg	Milligrams
MLM	Multilevel Model
MPQ	McGill Pain Questionnaire
MRI	Magnetic Resonance Imaging

NHP	Nottingham Health Profile
NHS	National Health Service
NOAA	National Oceanic and Atmospheric Association
NRS	Numerical Rating Scale
NSAIDs	Non-Steroidal Anti-Inflammatory Drugs
OA	Osteoarthritis
PBS	Pharmaceutical Benefits Scheme
PCR	Polymerase Chain Reaction
PESS	Patient Experienced Symptom State
PROs	Patient-Reported Outcomes
R^2	R-Squared Statistics
RA	Rheumatoid Arthritis
RAH	Royal Adelaide Hospital
RAID.7	Seven Items of Rheumatoid Arthritis Impact of Disease
RAPS	Rheumatoid Arthritis Pain Scale
RCT	Randomised Controlled Trial
RF	Rheumatoid Factor
RMDs	Rheumatic and Musculoskeletal Diseases
RPS	Regional Pain Scale
SC	Subcutaneous
sd	Standard Deviation
SDAI	Simplified Disease Activity Index
SES	Socioeconomic Status
SF-36	Short Form-36
SF-MPQ	Short-Form McGill Pain Questionnaire
SIJs	Sacroiliac Joints
SIP	Sickness Impact Profile
SJC	Swollen Joint Count

SMR	Standardised Mortality Ratio
SpA	Spondyloarthritis
T2T	Treat-to-Target
TJC	Tender Joint Count
TNF-α	Tumour Necrosis Factor-Alpha
TPS	Thermometer Pain Scale
UK	United Kingdom
VAS	Visual Analogue Score
VAS-GH	Visual Analogue Score-Global Health
VPN	Virtual Private Network
VRS	Verbal Rating Scale
WHYMPI	West Haven-Yale Multidimensional Pain Inventory
WOMAC	Western Ontario McMaster Questionnaire

Abstract

Extended Abstract

Problem Statement:

Chronic pain is a common sequela of rheumatic and musculoskeletal diseases (RMDs) worldwide. The health care and societal costs related to the burden of chronic pain are insurmountable. Chronic pain is regarded as a symptom of utmost importance to individuals living with arthritis. Chronic pain, to everyone, is unique, complex, and multidimensional, and for many, it can adversely impact on the individual's day-to-day physical functioning and psychosocial states.

In a typical rheumatology clinic, and usually with a standard 3-month or 6-month follow-up timeframe, patients are often asked to discuss their pain relating to their arthritis since the last visit, which is highly prone to recall bias and selective memory. The lived experience of pain in arthritis is highly variable and can be unpredictable at times, especially during disease flare or with other co-existing pain-related comorbidities, such as fibromyalgia (FM). Summarising the ebb and flow of pain may not necessarily reflect the real-time pain impact. Capturing these pain symptoms in real-time may provide the window of opportunity to intervene and to help better manage their pain when symptomatic. Furthermore, discordance between the retrospective summary of pain experience and the actual real-time impact of pain can result in unintended consequences of unnecessary treatment escalation and poor pain management. Data collection of repeated measures of symptoms and co-occurring events captured over time is well established, although it is often an onerous task and can be unappealing and intrusive. The implementation of mobile health (mHealth) in clinical practice and research has unfolded many potentials to capture temporally rich patient-reported outcomes (PROs) in real-time, when compared to traditional methods of data collection. In the context of using smartphones in capturing real-time pain symptoms in arthritis, this type of mHealth data creates a unique platform to

explore novel methods in examining pain trajectory and pain variability in RMDs. Such granular real-time data may provide opportunities to explore the key components of capturing pain 'flare', often defined as the momentary state of heightened/significant pain level and is usually accompanied by other pain-related symptoms.

Traditionally, pain in arthritis is understood as a form of nociceptive pain signalling output in the nervous system, often attributed to the disease activity such as synovial joint inflammation. Yet, despite adequately treated arthritis, PROs such as persistent pain and fatigue exist in some individuals, highlighting the roles of nociplastic central pain processing and the inter-relationship with psychosocial and socioeconomic factors. Despite treatment advances in disease modifying anti-rheumatic drugs (DMARDs) use in rheumatoid arthritis (RA), persistent pain remains a treatment conundrum to patients and clinicians, even in those with an absence of or low level of disease inflammation. Globally, the mortality gap in RA is high. Concerningly, chronic pain is known to increase mortality in the wider population, posing an important research question on whether mortality in RA is accelerated if the trajectory of health status is downtrending in those with persistent pain.

Research gap:

Following a systematic scoping review on the current literatures in arthritis, the phenomenon of persistent pain in arthritis is significantly associated with poorer quality of life and health outcomes, even in those with low disease activity. Heterogeneity in both the pain scales used and the frequency of pain measurements identified in these literatures informs the lack of standardised pain measurements and therefore, poses a challenge of applying intensive longitudinal methods in examining pain trajectory and variability in RMDs. Additionally, in RA, the association of persistent pain and its attributable burden on treatment and adverse health outcomes longitudinally is largely unknown.

Purpose statement:

By harnessing longitudinal data on pain-related PROs captured in large observational studies, understanding the trajectory of pain-related health outcomes in arthritis may provide new insights to patients, clinicians, and key stakeholders. Such research may improve the understanding of the phenomenon of persistent pain and its attributable burden on treatment and adverse health outcomes in arthritis, and more importantly, to identify those at risk and to intervene early with evidence-based strategies in managing their pain and their overall health and well-being.

This thesis aims to explore the longitudinal characteristics of pain in RMDs using intensive longitudinal methods and to determine the longitudinal effects of persistent pain on medication use and adverse health outcomes in arthritis such as RA.

Research questions:

1. What are the methods used to longitudinally assess pain in arthritis in the current literature?
2. What are the benefits and challenges in using digitalised health-related data in clinical care and research? In the context of harnessing patient-generated health data obtained from smartphones, what are the key outcomes that can be identified when assessing pain trajectory and variability using intensive longitudinal methods?
3. In RA, by using the validated pain score, disease activity composite score and other relevant PROs, can we identify individuals at risk of developing persistent pain trajectory? What is the impact of persistent pain on medication use and important health outcomes in RA, such as hospitalisation events, mortality risk and causes of death (COD)?

Methodology:

The research groundwork for this thesis commenced with my initial work being conducted in Manchester, United Kingdom (UK) in 2018 as a remote postgraduate student. My research work continued as I returned to South Australia in 2020 using Australian-based datasets. To address the research questions as outlined above, the research work is presented in 3 sections:

Section 1. Review of the methods used to examine longitudinal characteristics of pain in arthritis:

This systematic scoping review using 3 databases (Medline, Embase and Psycinfo) was conducted from inception to May 2022. Inclusion criteria for this review were of adults aged 18 years and older, with RA and with key components of longitudinal methodology applied in the included studies. Specifically, studies with repeated measures of pain of at least 3 timepoints or more were included. Data extraction included demographics of the included studies, baseline characteristics of the study cohorts, key components of the type of pain scales used and frequency of measurement and the relevant pain-related confounders or predictors associated with the primary analysis for each study. Key findings from these extracted data were critically discussed.

Section 2a. Review of the benefits and challenges in using digitalised health-related data in providing real-world evidence:

This narrative review outlined the opportunities of using digitalised health-related data and the challenges of data access, data quality and governance of data ownership, protection, and public trust. A case vignette of a patient with osteoarthritis (OA), from an asymptomatic state to disease onset and its trajectory, was presented. Key messages were outlined and discussed throughout this patient's health journey.

Section 2b: Analysis of pain trajectory and day-to-day pain variability in RMDs using a large, citizen scientist-led smartphone study data:

This work was conducted using *Cloudy with a Chance of Pain*, a national UK smartphone study. This citizen scientist-led smartphone study was conducted in Manchester, UK and study participants contributed their daily pain symptoms through a user-friendly mobile app. A total of 10,584 participants had complete data on baseline questionnaire and pain symptoms, with at least 2 or more timepoints of pain data entry. Specific for this thesis, a data subset of study participants with medically diagnosed inflammatory (RA and spondyloarthritis, SpA) and non-inflammatory (OA and FM) RMDs was included for the analysis. The overall pain trajectory and day-to-day pain variability for the first one-month period (pain level measured using ordinal 5-point scale: 1 – none, 2 – mild, 3 – moderate, 4 – severe, 5 – very severe) were analysed across these RMDs using multilevel and Markov transition models respectively. Day-to-day fluctuation in pain across these RMDs was also analysed.

Section 3a: Analysis of the DAS28-P index as a predictor of poor subjective response to DMARD therapy using an early RA cohort data:

DAS28 (28-joint Disease Activity Score) is a validated disease activity composite score used universally in RA, with the intention to guide treatment decision. The DAS28-P index, also known as the proportion of patient-reported components (joint tenderness and visual analogue scale (VAS)) of the DAS28, was identified as a relevant discriminatory measure of non-inflammatory pain in RA. Using an early RA cohort, derived from a supplemental fish oil clinical trial conducted at the Royal Adelaide Hospital, South Australia, consecutive key component measures of the DAS28-P index for each patient were obtained over a 1-year period. For the objective components of the DAS28 scores obtained in the study, apart from tender joint count, erythrocyte sedimentation rate (ESR) was used, instead of C-reactive protein (CRP). First, bivariate k-means model, a form of clustering analysis, was used to identify distinct subgroups of patients using these repeated measures of DAS28-P indices over time. Baseline predictors for each subgroup were assessed. Longitudinal outcomes of

disease activity, disease impact, joint erosion, and medication use were assessed using a random intercept, population-averaged generalised estimating equation model, which considered the variability between subgroups.

Section 3b: Analysis of pain-related health outcomes and longitudinal effects on medication use, sociodemographic indicators, hospitalisations, and mortality risk in RA using a large observational registry cohort data:

Study participants with rheumatologist-diagnosed RA in the Australian Rheumatology Association Database (ARAD), a national registry of participants with inflammatory arthritis contributing data on treatment safety and PROs, were included in this analysis. First, distinct multi-trajectory analysis using discrete mixture model was performed on 15-year follow-up data for five different self-reported pain-related health outcomes. These outcomes were Health Assessment Questionnaire Disability Index (HAQ-DI), VAS for pain, arthritis, global health, and the Assessment of Quality of Life (AQoL). With these identified trajectory subgroups of pain-related health status modelled as predictors, the associations with baseline demographics, sociodemographic indicators, and comorbidities. Medication use (opioids, non-steroidal anti-inflammatory drugs (NSAIDs), prednisolone and conventional synthetic DMARDs) was analysed using longitudinal panel mixed model regression. In terms of examining biologic DMARD modification, time-to-event analysis was performed, in which failure times in the model were defined at the initiation of, or change in, biologic DMARD. Multiple 'failures' (modifications of biologic therapy) were possible for each individual, necessitating the use of a random effects, parametric Weibull 'survival' model, which examined the baseline hazard rate and allowed for within-individual dependencies between treatment failure episodes. In addition, using similar predictors, hospitalisation events were examined using longitudinal logistic regression and mortality risk was calculated in the survival analysis which included Cox regression, with comparison matched to the Australian population mortality rates over 18 years of follow-up. Competing risk regression was used to examine International Classification of Diseases 10th Revision (ICD10)-defined COD in these trajectory subgroups.

Main findings:

Section 1: Longitudinal characteristics of pain in RA

In total, 22 out of 1,400 studies in the literature search were included and systematically reviewed in this scoping review. Overall, heterogeneity in pain scales used and the frequency of pain measurements was observed. Stable pain trajectories in RA were seen with minimal pain fluctuation over time. Two included studies identified distinct pain trajectory subgroups using clustering analysis and another two studies explored pain variability using novel dynamic time-series modelling approach, albeit short study period. Important predictors such as female gender, ethnicity, smoking history, and education level were observed in these included studies.

Section 2b: Pain trajectory and day-to-day pain variability in inflammatory and non-inflammatory RMDs

In this exploratory analysis, 1,189 participants with inflammatory (RA and SpA) and non-inflammatory (OA and FM) contributed 23,470 daily pain scores over the first one-month follow-up period (83% female; mean age of 50 years, median entry days of 26 [interquartile range (IQR) 21, 30]). The mean pain scores were higher in those with SpA and OA than in RA (2.73±0.98, 2.60±0.95 and 2.50±0.98 respectively; 57.4%, 51.3% and 45.8% reported moderate-very severe pain respectively). Participants with FM had the highest mean pain score of 3.04±1.03 (69.7% with moderate-very severe pain). In the multilevel model, the long-term change in pain was significantly different between participants, with steeper time-based improvements in pain for participants with higher initial pain scores across all diseases. In the Markov transition state model, highest probability of staying in the same pain states from yesterday was seen across all conditions (RA, SpA, OA, and CWP/FM). Similarly, from the day-to-day pain fluctuation plots, there was minimal absolute change in pain level from yesterday, defined as 2-point or 4-point increase and decrease in pain. 56% of participants with CWP/FM in the ‘very-severe’ pain state remained in this same pain state from yesterday and had minimal change in day-to-day pain state.

Section 3a: DAS28-P index as a discriminatory measure of treatment response in early RA

From 121 study participants (74% female; mean age of 57 years, median of 16 weeks of active disease, 54% seropositive for anti-cyclic citrullinated peptide antibodies (ACPAs)) with early RA on DMARDs, using the cluster analysis, 3 distinct subgroups were identified. These included 58 participants in the 'Responders' group (48%), 32 in the 'Partial Responders' group (26%), and 31 in the 'Non-Responders' group (26%). Baseline mean DAS28-ESR was 5.7 (standard deviation (sd) of 1.2). Throughout the study period, the 'Partial Responders' group remained having higher proportions of the DAS28-P index and had the lowest joint erosion score of 0.9 (95% confidence interval (CI) of 0.2, 1.6) when compared to the "Responders' (joint erosion score of 1.8, 95%CI 0.8, 2.7) and 'Non-Responders' (3.4, 95%CI 1.7, 5.2) groups. In terms of disease impact, the 'Responders' group had lower levels of function limitation, fatigue, and coping difficulty. At the end period of the study, both 'Partial Responders' and 'Non-Responders' groups used higher dose of methotrexate (18.5 milligrams (mg) [95%CI 15.5, 21.5] and 18.6mg [95%CI 15.3, 21.8] respectively) when compared with the 'Responders' group (12.8mg [95%CI 14.7, 20.9]). The median cumulative glucocorticoid dose was significantly higher in the 'Non-Responders' group (297mg of prednisolone equivalent, IQR 211, 284).

Section 3b: Trajectories of pain-related health status in RA and the associations with sociodemographic factors and medication use (Analysis Part 1) and hospitalisations and mortality risk (Analysis Part 2)

In Analysis Part 1, 988 ARAD study participants with RA (71% female; mean age of 54 years, mean disease duration of 2.3 years) were included. From the discrete mixture model, four, approximately equally sized, pain-related health status groups were identified, ranging from 'better' to 'poorer' groups. Changes over time between these trajectory groups were relatively minimal. In the poorer pain-related health status group, significant baseline predictors included female gender, obesity, smoking, sociodemographic indicators (lower education level, use of disability support, and lower socioeconomic status (SES)) and higher

comorbidity index. During follow-up, although the patterns of biologic use was similar between groups, biologic therapy modifications (p_{linear} <0.001) and greater preference of using non-tumour necrosis factor inhibitor (an increase in marginal probability from 0.09 to 0.21, p_{linear} <0.001) were seen in those with poorer pain-related health status group. Concerningly, in the time-varying longitudinal mixed model regression, the marginal probability of opioid use markedly increased from 0.08 in the 'better pain-related health status' group to 0.57 in the 'poorer pain-related health status' group. Prednisolone use was relatively high across all trajectory groups, with an increase seen in the marginal probability from 0.26 to 0.53 (p_{linear} <0.001).

In Analysis Part 2, 806 ARAD study participants with RA (72% female, mean age at diagnosis of 54 years) were included. Using similar trajectory analysis in Part 1, similar results were seen, with four, approximately equally sized, distinct pain-related health status trajectory groups were identified, ranging from 'better' to 'poorer'. In the poorer pain-related health status group, they were more likely to have obesity, higher comorbidity index and lower SES. In parallel, this group had the highest hospitalisation rates (odds ratio 3.09, 95%CI 2.3, 4.2). Excess mortality was seen in the poorer pain-related health status group, either when compared to the other three groups (hazard ratio 2.4, 95%CI 1.5, 3.8), or the matched Australian population (standardised mortality ratio (SMR) 2.3, 95%CI 1.6, 3.3, p<0.001). Circulatory diseases secondary to cardiovascular or cerebrovascular causes were the most frequently reported COD in this poorer pain-related health status group (subhazard ratio 8.3, 95%CI 2.2, 31.7).

Significance of findings and implications in clinical care and research:

Overarchingly, the work performed in this thesis imparts new knowledge of the application of intensive longitudinal methods to analyse pain trajectory and pain variability in RMDs. Additionally, specific to RA, these longitudinal methods have unravelled new insights of the attributable burden of persistent pain mediated by the longitudinal effects of unfavourable pain-related psychosocial constructs and poorer health status of those at-risk individuals.

The interpretations of the study findings and their implications in clinical care and research are outlined below:

- Pain in arthritis is multidimensional – the type of pain scales used and the frequency of pain measurements matter
- Pain trajectory analysis in arthritis allows stratification of individuals with different pain experience over time – at-risk individuals of having persistent pain are identifiable and early intervention could be addressed
- Changes in pain over time in arthritis is minimal, regardless of the frequency of pain measurements – alternative methods of assessing the magnitude of pain variability in real-time is required
- The role of the patient global assessment within the DAS28 in RA remains contentious when used to assess disease activity and to guide treatment – the DAS28-P index is shown to be a useful surrogate as the predictor of treatment response in RA
- In RA, pain is much more than just a symptom related to the arthritis itself – the pain experience and the health status of the individuals are interdependent
- In RA, a higher tendency to modify treatment or to increase treatment dose unnecessarily is seen in those with non-inflammatory pain and similarly, in those with worsening pain-related health status – in those individuals with persistent pain and poorer health status, a mindful evaluation of treatment decision and disease-related management is recommended at the time of diagnosis and throughout the disease course
- In RA, opioid use is shown to be markedly increased in those with worsening pain-related health status – at-risk individuals with persistent pain and using opioids should be identified and intervened early in the disease course to avoid opioid-related harms
- An increase in mortality risk in RA is a major concern in individuals with persistent pain, especially of those with poorer health status trajectory – a serious consideration of the influence of persistent pain on health outcomes in RA is just as

important as considering other well established risk factors that affect mortality in RA

Strengths and limitations:

The strengths of the research completed in this thesis are outlined below:

- The advantages of having research skills and opportunities gained early in the postgraduate study could be applied throughout the completion of this thesis – these include statistical programming skills, digital epidemiology, research skills in mobile health studies and the foundation in using intensive longitudinal methods
- A unique opportunity to harness different types of data sources to address the research questions, including the application of complex statistical methods to analyse different data sources, including the mobile health study and the large observational cohort studies

There are several limitations in the research as outlined below:

- In the work done using *Cloudy with a Chance of Pain* smartphone study, a short study period was selected during the analysis for data completeness. Hence, results from this exploratory work may be subject to selection bias, as highly engaged study participants were more likely to contribute data continuously.
- In the work done using the Australian based observational studies, study participants with RA were the focus in this thesis. Therefore, results from this work may not be generalisable to individuals with other forms of RMDs, such as OA and FM.
- Overall, the datasets used in the research for this thesis did not include information on disease activity, limiting further analysis of the relationship of the disease activity and the study outcome of interest such as pain-related health status. Furthermore, these datasets consist of mostly self-reported health-related information, which may be subject to recall bias or reporting bias.

Conclusion:

The study findings in this thesis have shown significant potentials that arise from the application of intensive longitudinal methods in analysing pain in arthritis. Pain trajectory analysis allows the opportunity to identify at-risk individuals, particularly of those with persistent pain. Changes in pain over time in arthritis is minimal, regardless of the frequency of pain measurements, suggesting the need for alternative methods when assessing pain volatility. The magnitude of the burden of persistent pain and poor health status in RA is significant, with concerning trends of higher tendency of unnecessary treatment use and modification and excess mortality risk. Most importantly, in the early days of the diagnosis of arthritis and throughout the disease course, pain and health status should be mindfully evaluated in parallel, with aims to improve the outlook of the individuals. In individuals with persistent non-inflammatory pain, the long-term adverse health-related impacts should not be an oversight amongst all treating rheumatologists and other health care professionals.

Acknowledgement

This thesis is a testament to the endless support and love I have received from many people and organisations throughout the years, and without them, I would not be where I am right now.

I would like to acknowledge the research funding support I received, both the Ken Muirden Overseas Training Fellowship from Arthritis Australia (an educational grant funded by the Australian Rheumatology Association and awarded in 2018) and the Australian Government Research Training Program Scholarship through the University of Adelaide.

To my primary supervisor, Professor Catherine Hill, I thank you from the bottom of my heart for your endless support, mentorship, and friendship, not only throughout my candidature, but from the start of my rheumatology career. My passion for research in rheumatology stems from your initial encouragement and support in gout-related collaborative research work, and subsequently, the overseas fellowship application and your unconditional mentorship and support throughout this candidature. There is no doubt that you are the precursor to the success of my research career in rheumatology.

To my external supervisor, Professor William Dixon, I thank you for the once-in-a-lifetime opportunity to be able to work with you and the team of the *Cloudy with a Chance of Pain* mobile health project in Manchester, UK. Thank you for allowing me to spend my first year of my candidature at the Centre for Epidemiology Versus Arthritis, University of Manchester and for introducing me to the world of digital epidemiology. I am always grateful for your time, patience, and guidance throughout my candidature.

To my local supervisors in Adelaide, I would like to thank Ms Sue Lester for your endless support in shaping and honing my research skills, especially in data analysis and interpretations. I am always learning something new from you, and I will always remember

our conversations on programming skills and our culinary and cooking talks outside of work. I am always grateful for your time and patience, particularly throughout the process of refining my research manuscripts and presentations. To Dr Sam Whittle, I thank you for your endless support and mentorship, not only throughout this candidature but also throughout my rheumatology career. Thank you, Sam, for giving me the opportunity to write and publish my first review paper on fibromyalgia. I am always grateful for your time, support, and feedback, especially with my research ideas and with refining my writing skills throughout my candidature, and for reading the final draft of this thesis. I thank Professor John Beltrame for your support and time throughout my candidature. I will always remember the support you gave me in my early years in medical school working on a summer vacation research project with you.

To my external supervisors in Manchester, UK, I thank Dr Mark Lunt and Professor John McBeth for your invaluable input and guidance, especially with my research work using the *Cloudy* data.

I thank the organisations and collaborators who have been pivotal in providing me with their support, time and invaluable feedback on the work included in this thesis. Specifically, in no particular order, I would like to thank Professor Susanna Proudman, Ms Leah McWilliams, Mr Rob Metcalf, Dr Rachel Black, Dr Joanna Tieu, Dr Oscar Russell, Professor Lyn March, Professor Rachelle Buchbinder, Professor Marissa Lassere, Dr Claire Barrett, Dr Bethan Richards, Mr Jaiden Nairne-Nagy, Dr Belay Birlie Yimer, Dr Anna Beukenhorst, Ms Louise Cook, Dr Thomas House, Dr Jamie Sergeant, Professor David Schultz and Dr Meghna Jani for their support.

Special mention to Rachel and Joanna, thank you for your time and friendships. I do miss the lunch breaks and the ad hoc conversations we had at the Basil Hetzel Institute. Also, I want to thank Belay and Anna for the intellectual and fun talks we had in the department, especially on number crunching, fixing data bug, and generating new codes in R. I would

not have been an R programmer without your support and patience, Belay and Anna, and I am forever grateful to you both. Special thanks to Meghna for being a great and supportive mentor and friend to me when I was in Manchester.

Thank you to the study participants and the study steering committee and staff members involved in the work I included in this thesis. Thank you for your contributions and hard work in these study projects – *Cloudy with a Chance of Pain*, Early Rheumatoid Arthritis Cohort in South Australia, and the Australian Rheumatology Association Database (ARAD).

Thank you to my Postgraduate Coordinators, Professor Betty Sallustio and Dr Peter Zalewski for your support and feedback throughout my candidature, especially with my annual review assessments.

I am forever grateful to my parents (Kok Beng and Siew Moy), my siblings (Huai Sun, Huai Peng and Huai Poh), and my in-laws (Dino, Lisa, Sandra and Adrian) for their endless love and encouragement in my career and life.

To my family, in the last few years of my candidature, I am still in awe that we have done so much – moving countries, surviving through the Covid-19 pandemic, and facing the ups and downs in life and careers. I am grateful that we have each other. To my beautiful daughters Miranda and Naomi, you are the love of my life, and I cannot wait to spend more time with you both. To my wonderful husband, Anthony, you are the pillar in our family, holding everything strong and steady, especially in the last few years – I love you dearly and thank you for everything.

Dedication

I dedicate this thesis to Anthony, Miranda, and Naomi. To Anthony for your endless love and support emotionally and financially. To Miranda and Naomi for the fun and crazy times – completing this PhD journey with you, although challenging, has taught me motherhood, hope, and faith in conquering the impossible.

Publications During Candidature

Publications contributing to this thesis

1. **Pisaniello, H.L.**, Dixon, W.G. What does digitalization hold for the creation of real-world evidence? Rheumatology (Oxford, England), 2020; 59(1):39-45. doi:10.1093/rheumatology/kez068 [**Published**]
2. **Pisaniello, H.L.**, Whittle, S.L., Lester, S., Menz, F., Metcalf, R., McWilliams, L., Hill, C.L., Proudman, S. Using the derived 28-joint disease activity score patient-reported components (DAS28-P) index as a discriminatory measure of response to disease-modifying anti-rheumatic drug therapy in early rheumatoid arthritis. BMC Rheumatology, 2022; 6:67. doi.org/10.1186/s41927-022-00299-3 [**Published**]
3. **Pisaniello, H.L.**, Lester, S., Russell, O., Black, R.J., Tieu, J., Richards, B., Barrett, C., Lassere, M., March, L., Buchbinder, R., Whittle, S.L., Hill, C.L. Trajectories of Self-Reported Pain-Related Health Outcomes and Longitudinal Effects on Medication Use in Rheumatoid Arthritis: A Prospective Cohort Analysis Using the Australian Rheumatology Association Database (ARAD) [**Published**]
4. **Pisaniello, H.L.**, Nairne-Nagy, J., Lester, S., Whittle, S.L., Hill, C.L. Longitudinal Characteristics of Pain in Rheumatoid Arthritis: A Scoping Review [**Submitted**]
5. **Pisaniello, H.L.**, Lester, S., Russell, O., Black, R.J., Tieu, J., Richards, B., Barrett, C., Lassere, M., March, L., Buchbinder, R., Hill, C.L., Whittle, S.L. The Associations Between Poorer Pain-Related Health Status and Increased Hospitalisations and Excess Mortality in Patients with Rheumatoid Arthritis: A Prospective Cohort Analysis Using the Australian Rheumatology Association Database (ARAD) [**Submitted**]

Other relevant publications

1. Dixon, W.G., Beukenhorst, A.L., Yimer, B.B., Cook, L., Gasparrini, A., El-Hay, T., Hellman, B., James, B., Vicedo-Cabrera, A.M., Maclure, M., Silva, R., Ainsworth, J., **Pisaniello, H.L.**, House, T., Lunt, M., Gamble, C., Sanders, C., Schultz, D.M., Sergeant, J.C., McBeth, J. How the weather affects the pain of citizen scientists using a smartphone app. NPJ Digital Medicine, 2019; 2:105. doi.org/10.1038/s41746-019-0180-3 [**Published**]
2. Schultz, D.M., Beukenhorst, A.L., Yimer, B.B., Cook, L., **Pisaniello, H.L.**, House, T., Gamble, C., Sergeant, J.C., McBeth, J., Dixon, W.G. Weather Patterns Associated with Pain in Chronic-Pain Sufferers. Bulletin of the American Meteorological Society (BAMS), 2020; 101(5). doi/org/10.1175/BAMS-D-19-0265.1 [**Published**]
3. Yimer, B.B., Schultz, D.M., Beukenhorst, D.M., Lunt, M., **Pisaniello H.L.**, House, T., Sergeant, J.C., McBeth, J., Dixon, W.G. Heterogeneity in the association between weather and pain severity among patients with chronic pain: a Bayesian multilevel regression analysis. Pain Reports, 2022;7doi.org/10.1097/PR9.0000000000000963 [**Published**]

Presentations During Candidature

1. **Pisaniello, H.L.** The role of mobile health application in real-time capture of self-reported symptoms and longitudinal activity, and its feasibility in patient-focused remote monitoring in musculoskeletal disorders. Major Review oral presentation at the Basil Hetzel Institute (BHI) Seminar, Woodville South, South Australia, February 2019.

2. **Pisaniello, H.L.** Opportunities from daily patient-generated data: results from the *Cloudy with a Chance of Pain* study. Oral presentation at the Centre for Epidemiology Versus Arthritis Seminar, Manchester, UK, June 2019.

3. **Pisaniello, H.L.** Examining long-term and short-term day-to-day pain variability using *Cloudy with a Chance of Pain* dataset. Oral presentation for the Daily Symptom Method meeting, Manchester, UK, April 2020.

4. **Pisaniello, H.L.** Navigating the tides and waves of chronic pain. 3-minute Thesis finalist oral presentation at the University of Adelaide, Adelaide, South Australia, August 2020.

5. **Pisaniello, H.L.**, Lunt M., Lester, S., Whittle, S., Hill, C., McBeth, J., Dixon, W. Long-term and short-term day-to-day pain variability in inflammatory and non-inflammatory rheumatic diseases: multilevel and Markov transition analyses of the *Cloudy with a Chance of Pain* study. New Investigator Award oral presentation for the Australian Rheumatology Association (ARA) Annual Scientific Meeting, August 2020. BHI Seminar Sept 2020.

6. **Pisaniello, H.L.**, Lunt M., Lester, S., Whittle, S., Hill, C., McBeth, J., Dixon, W. Analysing the long-term and short-term day-to-day pain variability in inflammatory and non-inflammatory rheumatic diseases using multilevel and Markov transition models: *Cloudy with a Chance of Pain*, a national UK smartphone study. Virtual poster presentation for the Florey Postgraduate Research Conference, the University of Adelaide, Adelaide, South Australia, August 2020.

7. **Pisaniello, H.L.**, Lunt M., Lester, S., Whittle, S., Hill, C., McBeth, J., Dixon, W. Day-to-day pain variability in rheumatic and musculoskeletal diseases: *Cloudy with a Chance of Pain*. Oral presentation for the BHI Seminar, Woodville South, South Australia, September 2020.

8. **Pisaniello, H.L.**, Lunt M., Lester, S., Whittle, S., Hill, C., McBeth, J., Dixon, W. Examining day-to-day pain variability in inflammatory and non-inflammatory rheumatic diseases using multilevel and Markov transition models: *Cloudy with a Chance of Pain*, a nationwide UK smartphone study. Oral presentation for the Queen Elizabeth Hospital Research Expo, Woodville South, South Australia, October 2020.

9. **Pisaniello, H.L.**, Lunt M., Lester, S., Whittle, S., Hill, C., McBeth, J., Dixon, W. Analysing day-to-day pain variability in inflammatory and non-inflammatory rheumatic diseases using multilevel and Markov transition models: *Cloudy with a Chance of Pain*. Oral presentation for the ARA (South Australia branch) state annual scientific meeting, Adelaide, South Australia, October 2020.

10. **Pisaniello, H.L.**, Lunt M., McBeth, J., Dixon, W. Examining the long-term and short-term day-to-day pain variability in inflammatory and non-inflammatory rheumatic and musculoskeletal diseases using multilevel and Markov transition models: *Cloudy with a Chance of Pain*, a nationwide UK smartphone study. Virtual poster presentation for the American College of Rheumatology Convergence Scientific Meeting, USA, October 2020.

11. **Pisaniello, H.L.**, Lunt M., Lester, S., Whittle, S., Hill, C., McBeth, J., Dixon, W. Examining day-to-day pain variability in inflammatory and non-inflammatory rheumatic diseases using multilevel and Markov transition models: *Cloudy with a Chance of Pain*, a nationwide UK smartphone study. Oral presentation at the Centre for Epidemiology Versus Arthritis Seminar, Manchester, UK, January 2021.

12. **Pisaniello, H.L.**, Lunt M., Lester, S., Whittle, S., Hill, C., McBeth, J., Dixon, W. Examining day-to-day pain variability in rheumatoid arthritis – review of methods, challenges, and future directions: *Cloudy with a Chance of Pain*, a nationwide UK smartphone study. Oral presentation for the ARA (South Australia branch) state annual scientific meeting, Adelaide, South Australia, October 2021.

13. **Pisaniello, H.L.** Examining day-to-day pain flare in rheumatic and musculoskeletal diseases: Cloudy with a Chance of Pain. Oral presentation for the Cumulative Nimbus meeting, Centre for Epidemiology Versus Arthritis, University of Manchester, UK, November 2021.

14. **Pisaniello, H.L.**, Lester, S., Russell, O., Black, R., Tieu, J., Richards, B., Barrett, C., Lassere, M., March, L., Buchbinder, R., Whittle, S., Hill, C. Trajectories of self-reported pain-related health outcomes and the association with sociodemographic factors and medication use in rheumatoid arthritis: a prospective cohort analysis using the Australian Rheumatology Association Database (ARAD). Poster tour oral presentation for the ARA Annual Scientific Meeting, Hobart, Tasmania, May 2023.

15. **Pisaniello, H.L.**, Lester, S., Russell, O., Black, R., Tieu, J., Richards, B., Barrett, C., Lassere, M., March, L., Buchbinder, R., Hill, C., Whittle, S. Poorer pain-related health status in patients with rheumatoid arthritis correlates with an increased in hospitalisations and mortality: a prospective cohort analysis using the ARAD. Poster tour oral presentation for the ARA Annual Scientific Meeting, Hobart, Tasmania, May 2023.

16. Pisaniello, H.L. Trajectories of self-reported pain-related health outcomes in rheumatoid arthritis (ARAD). Oral presentation for the Australian Arthritis and Autoimmune Biobank Collaborative (A3BC)-ARAD Collaborative meeting, Sydney, New South Wales, October 2023.

Scholarships and Awards

Date	Type	Title	Institution Name
2018	Scholarship	Ken Muirden Overseas Fellowship	Arthritis Australia
2018 – 2020	Scholarship	Australian Government Research Training Program Scholarship	The University of Adelaide
2020	Award	3-minute thesis finalist	The University of Adelaide
2020	Award	New Investigator Award for the ARA Annual Scientific Meeting	The Australian Rheumatology Association
2023	Award	NHMRC CRE Better Outcomes in Inflammatory Arthritis Award for Best Completed Project	The Australian Arthritis and Autoimmune Biobank Collaborative (A3BC) – Australian Rheumatology Association Database (ARAD)

Chapter 1: Introduction

1.1 Preface

Each human is no stranger to pain. For a long time, pain is not only known as a form of physical discomfort in the medical field, but it is also ubiquitous in the society from the philosophical, religious, and political perspectives. Most individuals living with rheumatic and musculoskeletal diseases (RMDs) have experienced pain, either acute or chronic, early in the diagnosis and throughout the disease course. In arthritis, besides the well-established pathophysiological understanding of pain, recognising the ebb and flow of pain experience from the patient's perspective is equally important in the management of pain and the underlying disease. On this account, the approach on how we evaluate pain and the longer-term impact of pain on individuals living with arthritis forms the basis of this thesis.

This chapter provides an overview of pain in arthritis from a pathophysiological perspective and the role of central pain processing in persistent pain. In addition, this chapter summarises the need for longitudinal assessment of pain in arthritis and specifically, on potential analytical approaches when assessing longitudinally tracked data on pain symptoms derived from mobile health (mHealth) studies and observational studies. The attributable burden of persistent pain on treatment and health outcomes in arthritis is also discussed. This chapter then concludes by addressing the thesis aims and outline.

1.2 Background

"Pain has an element of blank;
It cannot recollect
When it began, or if there were
A day when it was not.

It has no future but itself,
Its infinite realms contain
Its past, enlightened to perceive
New periods of pain."

Emily Dickinson (1890)[1]

This poem, written in the 19th century by a famous American poet, Emily Dickinson, encapsulates the overall phenomenology of pain[1]. She eloquently described pain in its nature, temporality, and the totality of its existence through time in individuals living with chronic pain[2]. Historically, pain has been described ever since the existence of human beings, and over time, the concept of pain has evolved from the pre-historic era through to the current modern world[3]. Inherently, pain perception is subjective and unique to everyone.

In most RMDs, pain is one of the predominant clinical features, yet the pain perceived can differ between individuals. Additionally, within the same individual, pain in arthritis varies over time and can have different day-to-day impact on the person's physical function and psychological health. The next section explores the topic of chronic pain in arthritis. An overview of these RMDs – rheumatoid arthritis (RA), osteoarthritis (OA), spondyloarthritis (SpA), and fibromyalgia (FM) – is first presented, of which the purpose is to introduce these conditions for the readers of my thesis. Detailed discussion of these RMDs is beyond the scope of my thesis. Following the overview of these RMDs, I present a focused discussion on the pathogenesis and the role of neuroinflammatory pathways of chronic pain in arthritis.

1.3 Chronic Pain in Arthritis

Arthritis is a form of disease that affects the joints and associated peri-articular structures. For patients living with inflammatory and non-inflammatory arthritis, to them, pain is considered a symptom of utmost importance[4]. Their history of pain, recent or past, is commonly discussed between the patients and the clinicians.

In a typical clinic consultation, when patients are asked about their pain history, this is often done through a scramble of memory recollection that sums up their overall pain experience over a pre-specified timeframe. Such retrospective history taking may subject the patients to recall only the worst pain experienced, or to some, pain can be reported as being stable with no change. In managing arthritis, this pertinent information provided by the patients is heavily relied upon by the treating clinicians to justify the treatment decision and management of their pain and arthritis. However, such recall of the pain experience may not necessarily reflect the actual ebb and flow of the patient's pain experience, and in a mismatch between disease activity and pain, the pain mechanisms are not simply explained by the underlying arthritis itself.

In arthritis, pain is classically attributed to the disease activity such as joint inflammation, also known as the peripheral nociceptive pain processing. However, for some patients, despite having well-controlled disease, they may experience persistent pain, often due to central pain pathways[5-7]. Pain can also differ in the course of time along with the disease progression and treatment. Over time, the lived experience of pain is invariably influenced by the trajectory of the individual's health status and other biopsychosocial factors. Overall, chronic pain in arthritis is complex and multifaceted. This remains an ongoing research conundrum on how we can better evaluate the complexity of pain, considering the pathophysiology of pain but also the influence of biopsychosocial factors and general health status and wellbeing on the pain trajectory. What follows is a brief overview of the RMDs of interest.

1.3.1 Rheumatoid arthritis (RA) – an overview

RA is a chronic autoimmune disease, which commonly results in painful and swollen joints, and can present in other peri-articular and extra-articular systemic manifestations[8-10]. Globally, the prevalence of affected adults with RA is 5 per 1000[8]. In Australia, according to the latest 2017-2018 Australian Bureau of Statistics (ABS), nearly half a million of adult Australians are diagnosed with RA, representing 1.9% of the total population and 13% of all RMDs in Australia[11]. RA affects predominantly women, with at least 2 to 3 times more than men, and can occur at any age, with a peak incidence between the age of 50 to 60 years[8, 10].

Although the cause of RA remains unknown, the antigen presenting gene, *HLA-DRB1* and cigarette smoking are some of the well-established genetic and environmental risk factors for developing RA[12-17] (Figure 1.1). Autoantibodies such as rheumatoid factor (RF) and anti-citrullinated protein antibodies (ACPA) are commonly identified in patients with active RA (termed as seropositive), due to the epigenetic alterations and breakdown of immune tolerance in the immune system and subsequent cascades of synovial inflammation[8, 18]. These autoantibodies can be absent in one third of patients with RA (termed as seronegative) and can exist long before the appearance of RA symptoms (also known as pre-RA)[8-10, 19].

In the early stage of RA, the clinical features usually range from oligoarticular to severe polyarticular involvement in a symmetrical distribution and with minimal structural damage[8, 20]. If left untreated or inadequately treated, up to 90% of patients with RA can progress to more severe, erosive articular disease and other systemic or extra-articular features may develop[21]. Such progressive and destructive clinical features are uncommon nowadays as individuals with suspected RA are diagnosed and treated early by implementing the 'treat-

to-target' (T2T) approach and the use of disease-modifying anti-rheumatic drugs (DMARDs).

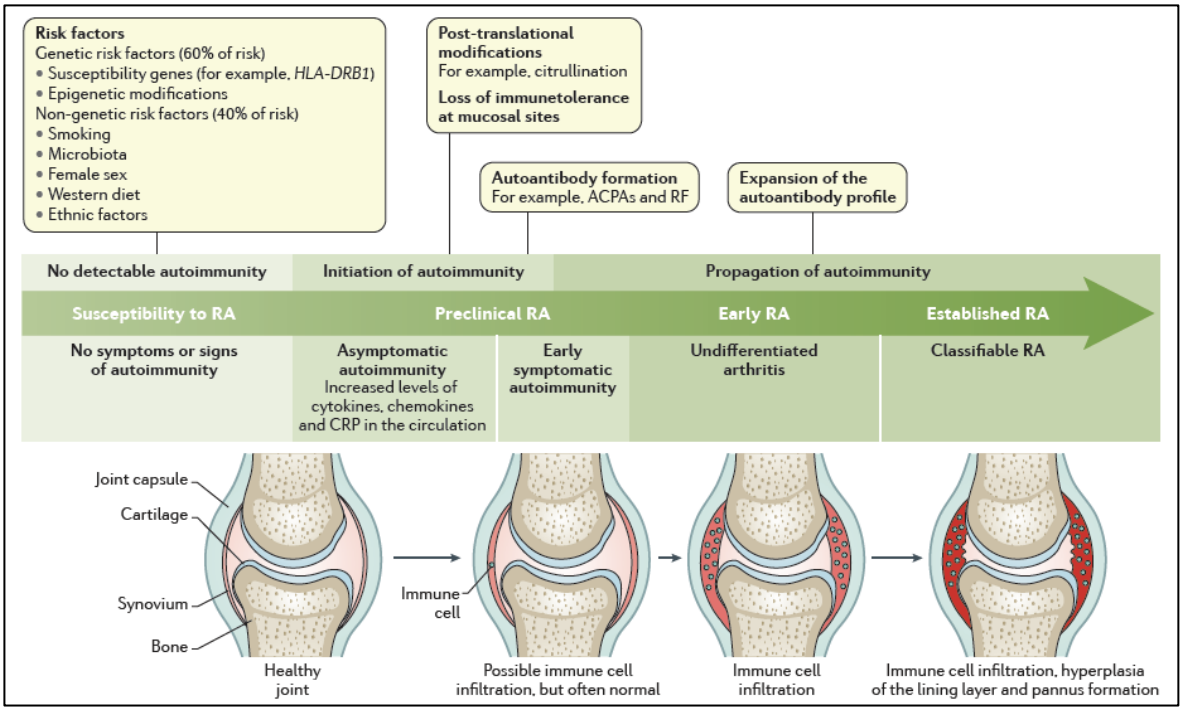


Figure 1.1 Risk factors and disease progression in rheumatoid arthritis

Source: Adapted from Smolen JS, Aletaha D, Barton A, et al[9] (used with permission under the Springer Nature – Copyright Clearance Center's Rightslink)

In 2010, the American College of Rheumatology (ACR) and the European Alliance of Associations for Rheumatology (EULAR; formerly known as the European League Against Rheumatism) collaboratively redefined the RA classification criteria, which significantly changed the paradigm of how we treat RA[22]. With the primary goal of early identification and intensive treatment early on in RA, this 2010 RA classification criteria set provides a standardised approach in identifying homogenous cohort of patients for clinical studies and informs clinical diagnosis and teaching[22-24].

Swollen and painful small joints involving the hands (metacarpophalangeal and proximal interphalangeal joints), wrists, knees, and feet (metatarsophalangeal joints), prolonged morning stiffness (≥30 minutes) and positive autoantibodies are typically seen in those with suspected RA[25]. Apart from clinical examinations and serological markers, composite

measures, such as the 28-joint disease activity score (DAS28), the clinical disease activity index (CDAI) or the simplified disease activity index (SDAI), are used to assess disease activity and to correlate with radiologic progression and functional outcomes[26, 27]. The degree of RA disease activity can be summarised and stratified using the specific cut points of these composite indices, which then facilitate treatment decision to achieve low disease activity or remission[28, 29]. DMARDs, classified as conventional synthetic (csDMARDs) and biologic/targeted synthetic (b/tsDMARDs) (Table 1.1), have revolutionised not only in improving disease suppression in RA, but also in preventing RA-related complications such as irreversible joint destruction and cardiovascular events and RA-related disabilities in terms of physical function and health-related quality of life[8].

Conventional synthetic DMARDs	Methotrexate, sulfasalazine, leflunomide, hydroxychloroquine
Biologic/targeted synthetic DMARDs	Etanercept, infliximab, adalimumab, golimumab, certolizumab, tocilizumab, abatacept, rituximab, tofacitinib, baricitinib, upadacitinib

Table 1.1 Currently approved disease-modifying anti-rheumatic drugs (DMARDs) for rheumatoid arthritis in Australia

1.3.2 Osteoarthritis (OA) – an overview

Synovium is the scaffold of human joints, which comprises connective tissue that lines the joint and encapsulates the tendons and the sheaths of the bursae and fat pads surrounding the joints. In OA, dynamic structural changes occur not only in the synovium, but also in the subchondral bone, hyaline articular cartilage, ligaments, and peri-articular muscles, and therefore, OA is also known as a disease of the whole joint[30-33]. Globally, OA is considered the most common arthritis, commonly affecting the knee, hip, and hand joints, and is the leading cause of arthritis-related pain and disability[31, 34-36]. Based on the 2017-18 ABS, at least 2.2 million adult Australians have OA, representing 62% of all RMDs in Australia[37]. One in five Australians over the age of 45 years suffer from OA[37]. OA predominantly affects women and can occur at all ages, although the prevalence is much higher from the age of 45 years[31, 36, 37]. Broadly, age is the strongest risk factor for OA development[31, 38]. In knee

and hip OA, the strongest risk factors include heavy workload in farming and construction sectors, repetitive loading in the joints, and certain high-impact and contact sports[31, 39-42]. Female gender, high BMI, and previous joint injury are associated with higher risk of developing knee OA, although less so for hip OA[31, 38, 43]. Clinical diagnosis of OA is mainly based on presenting complaints of joint pain, which is usually made worse with activity, minimal morning stiffness and limitation in physical function, as well as clinical signs of restricted range of motion and associated crepitus in the affected joint(s). Imaging and laboratory testing are not required to confirm the diagnosis. Education and self-management, as well as physical activity, and weight management remain the current recommended treatment strategies for OA[31, 44, 45]. Paracetamol and non-steroidal anti-inflammatory drugs (NSAIDs) can be used in managing painful OA[31, 45]. Effective DMARDs for OA are yet to be available for use[31]. Elective joint procedures such as arthroscopy and the use of opioids have minimal benefits in managing symptomatic OA. Careful evaluation of suitable candidates for surgical joint replacement in symptomatic patients with advanced OA is necessary for optimising outcomes.

1.3.3 Spondyloarthritis – an overview

Historically, individuals with inflammatory back pain and evidence of radiographic inflammatory changes in the thoracolumbar spine and sacroiliac joints (SIJs) were described as having ankylosing spondylitis[46, 47]. However, this term 'ankylosing spondylitis' does not cover the entity of early inflammatory phase in these affected axial regions long before the development of joint erosion, of which disease control is pertinent at this early, active inflammatory phase[46]. Nowadays, the term 'spondyloarthritis' (SpA) is a recommended descriptive term which encompasses a spectrum of inflammatory disease that affects primarily the spine (axial spondyloarthritis, including sacroiliitis), and may involve other human organs. These include the peripheral joints (peripheral inflammatory arthritis), entheseal structures (peripheral enthesitis), swollen digits (dactylitis), eyes (uveitis, conjunctivitis), skin (psoriasis), urinary system (urethritis), gastrointestinal system (inflammatory bowel disease), and large vessels (aortitis)[46] (Figure 1.2).

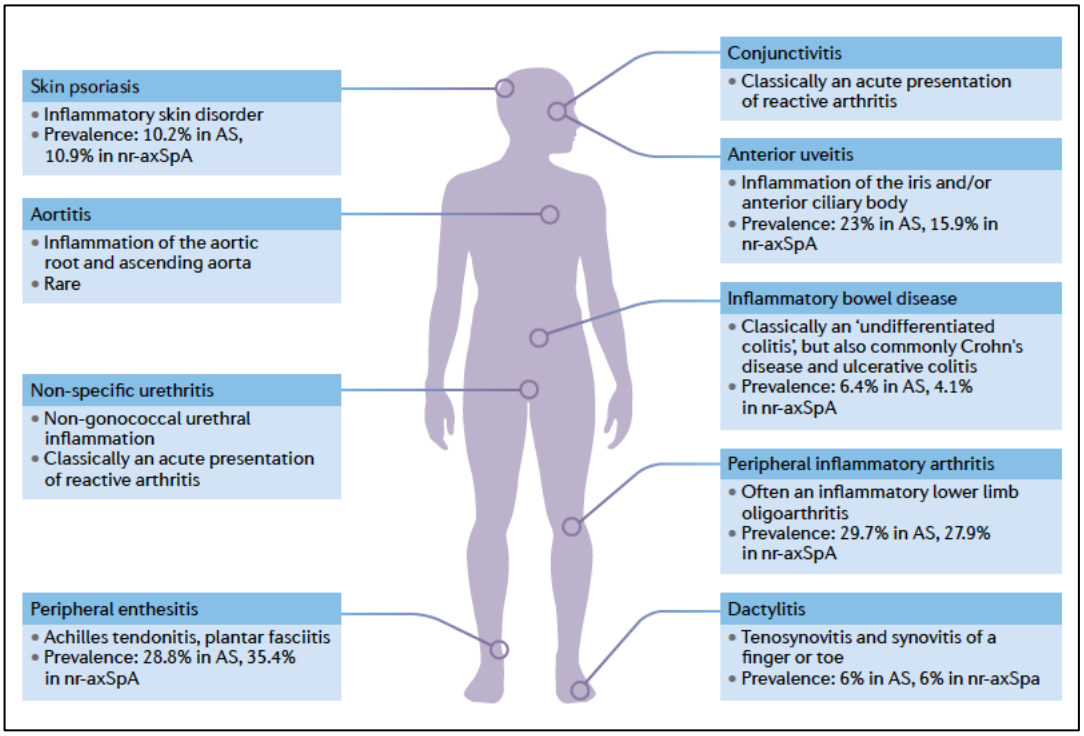


Figure 1.2 Clinical features of spondyloarthritis

Source: Adapted from Robinson, PC, van der Linden S, Khan MA, et al[46] (used with permission under the Springer Nature – Copyright Clearance Center's Rightslink)

Plain X-ray of the SIJs, in the presence of apparent sacroiliitis, helps to confirm the diagnosis of axial SpA[46]. However, magnetic resonance imaging (MRI) can be helpful in identifying early changes such as bone marrow oedema or fatty lesions in the SIJs, if X-ray findings are inconclusive[46].

1.3.4 Fibromyalgia (FM) – an overview

Pain is the most significant symptom described by individuals suffering from FM, representing one of the most common causes of chronic widespread pain (CWP). However, apart from pain, other symptoms such as fatigue, sleep impairment and undifferentiated somatic symptoms (also termed as 'functional symptoms') are equally described in individuals with FM[48] (Figure 1.3).

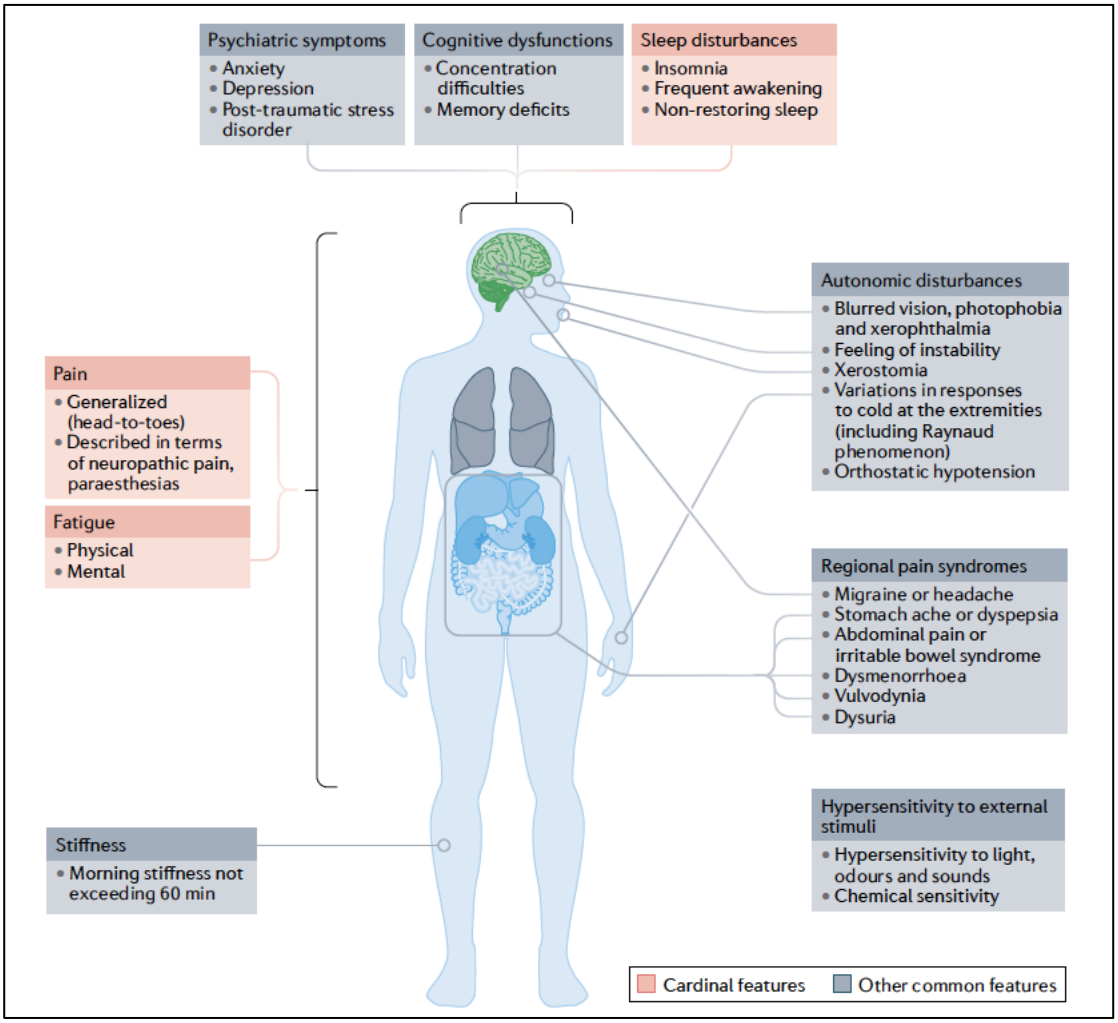


Figure 1.3 Clinical features of fibromyalgia

Source: Adapted from Sarzi-Puttini P, Giorgi V, Marotto D, et al[48] (used with permission under the Springer Nature – Copyright Clearance Center's Rightslink)

FM has a global prevalence of 2-3%, which is proportional to the age of those affected (peak age of onset at 50-60 years) and is more common in women (female:male ratio of 3:1)[48-50]. Diagnosing FM is unequivocally clinical, usually with detailed history-taking, and physical examination is not diagnostically supportive[48, 51]. The treatment of FM is based on a holistic, individualised, multi-modal biopsychosocial approach, as FM is considered as a condition with an overdriven mind-body connection in the context of genetic factors, personal lived experience, emotions and cognitive factors, and stress-coping mechanisms[48, 52-54]. Education, physical activity, psychotherapy, and effective pharmacological interventions are the recommended key treatments for FM[48].

In the next section, I will discuss the constructs of pain and inflammation, with RA and OA as the RMDs of interest highlighted in this topic, and specifically, pain in RA will be discussed in greater detail, as it features the most in my research work presented in this thesis.

1.3.5 The relationship between pain and inflammation

Historically, Celsus' definition of inflammation consists of four important clinical signs – erythema, heat, swelling and pain[55, 56]. Two centuries later, a fifth sign of loss of function was introduced, as Galen proposed his idea of the circulatory effects on inflammation[56, 57]. Nonetheless, this fifth cardinal sign of inflammation was only formally included by Virchow in the 19th century. Based on this definition, pain was originally considered as a localised output from the inflammatory process. Shortly after, in 1901, researchers discovered that inflammation and sensory neuron activation were interdependent and thereafter, the term 'neurogenic inflammation' was introduced[58]. Physiologically, the inflammatory mediators activate the primary afferent fibres (known as the nociceptors) and the sympathetic peripheral nerve fibres, resulting in a cascade of inflammatory features including pain, localised swelling, vasodilation, and increased warmth. In return, at the inflammatory site, these nociceptors activate local axonal reflexes, and subsequent neurotransmitters are released, stimulating ongoing neurogenic inflammation. These nociceptors also signal the pain pathway in the central nervous system (CNS), by activating the dorsal root reflexes in the spinal cord and the neuroendocrine and autonomic systems in the brain. Inflammatory pain, as we now understand, is a product of the homeostatic interaction between the inflammatory process and the nervous system. This interaction can become dysregulated in chronic inflammation, as seen in arthritis.

Morphologically, as shown in Figure 1.4, a healthy joint is innervated throughout with nociceptors and sympathetic nerve endings, and the innervation is more widespread in an OA joint[59-62].

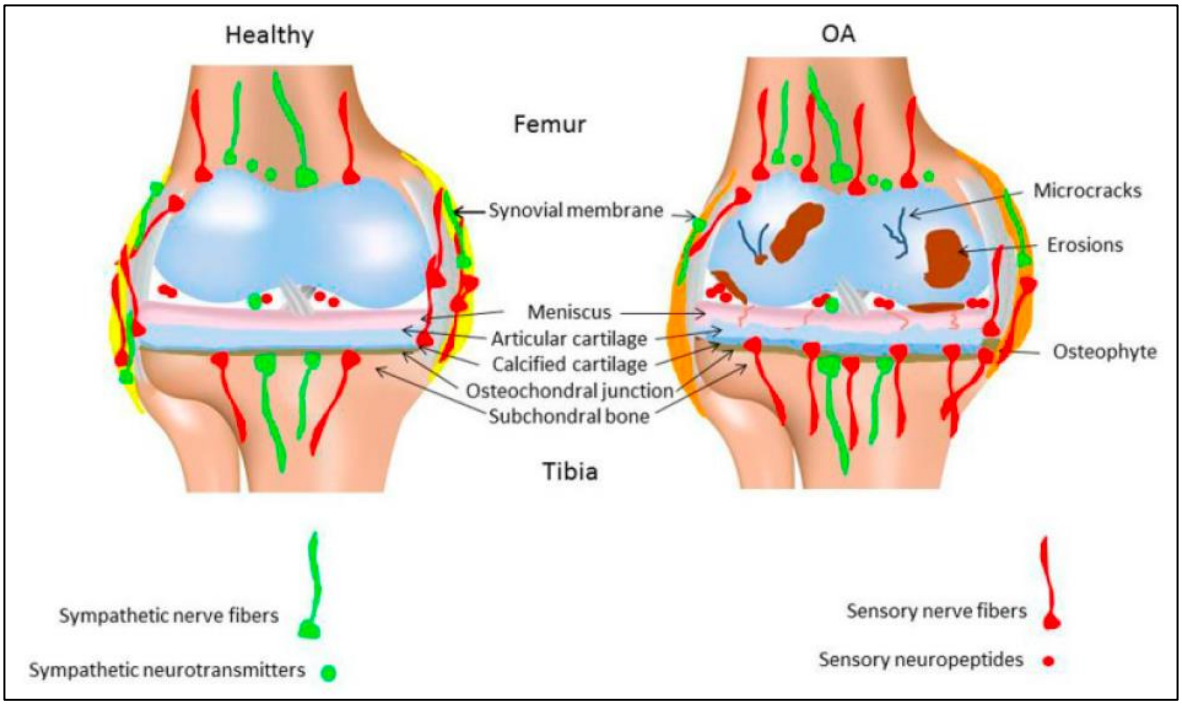


Figure 1.4 An illustration of the key components of sensory and sympathetic nervous system compared between healthy and osteoarthritic knee joints

Source: Adapted from Grassel S and Muschter D[63] (used with permission under Creative Commons Attribution 4.0 International)

These nerve fibres are found in the ligaments, subchondral bone, periosteum, synovium, and the outer third of the menisci, but not the cartilage[59-62]. Myelinated Aβ (with fast conduction) and unmyelinated C (with slow conduction) are the primary sensory nerve fibres involved in response to different stimuli. Specific to the joints, the Aβ nerve fibres are activated by the movement of the articular structures, and both Aβ and C fibres are activated by noxious stimuli, such as injury or inflammation. In inflamed joints, there is disruption to the osteochondral junctions, with erosive changes and structural remodelling process occurring in RA and OA respectively. Notably, there is also an increase in the number of nerve fibres and neurotransmitters in these osteochondral junctions, primarily due to angiogenesis and subsequent development of new nerve fibres[60, 64, 65]. In the context of inflammation, these nerve fibres can be activated at a lower threshold in response to the stimuli, known as sensitisation. This phenomenon is crucial in understanding the development of pathological pain state in arthritis, which is reviewed next.

1.3.4 The mechanisms of pain sensitisation in arthritis

In an inflamed joint, the afferent nerve fibres are sensitised through an array of pro-inflammatory mediators, as outlined in Table 1.2[64, 66-70].

Mediators at nociceptive endings	Cytokines, proteases, neuropeptides, chemokines, prostaglandins, neurotrophins, gaseous mediators, lipids
Mediators at dorsal root ganglion	Nerve growth factor, calcitonin gene-related peptide, vasoactive intestinal peptide, vanilloid receptor 1, opioid receptors, CC-chemokine ligand 2, CC-chemokine ligand 2 receptor
Mediators at brain level	Substance P, serotonin, glutamate

Table 1.2 Mediators of pain and inflammation in arthritis

Source: Adapted and restructured from Fu K, Robbins SR and McDougall JJ[64] (used with permission under the Oxford University Press – Copyright Clearance Center's Rightslink)

In return, there is higher propensity for these nociceptors to be activated in response to the noxious and mechanical stimuli[5, 60]. As it occurs in the peripheral nerve endings, this process is known as the peripheral sensitisation.

In a sustained inflamed joint, the CNS plays an important role, with the neurons in the spinal cord being activated secondary to the constant hyperexcitability of the peripheral nociceptors. As RA consists of a systemic inflammatory process, circulatory cytokines can potentially cross the blood-brain barrier within the CNS[68, 71]. In addition, internal cytokines such as interleukin-1 (IL-1) are generated from the immune cells in the CNS, further perpetuating the process of central pain sensitisation[68, 72, 73]. Peripherally and centrally, there is an increased responsiveness to noxious stimuli (also known as hyperalgesia) and to normally innocuous stimuli (also known as allodynia), as summarised in Figure 1.5[5, 64, 68].

Synovitis in RA and inflammatory OA augments the central pain processing. Both hyperalgesia and allodynia develop, due to continuous exposure to pro-inflammatory

cytokines such as tumour necrosis factor-alpha (TNF-α), IL-1 and interleukin-6 (IL-6) within the CNS[68, 74]. Interestingly, some animal studies have shown the involvement of central pain processing with specific histopathological, clinical characteristic and electrophysiological changes, pre-dating clinical evidence of active synovitis[68, 72, 75]. These findings highlight the importance of central pain pathways in chronic pain in arthritis, particularly in the realm of persistent pain despite disease suppression.

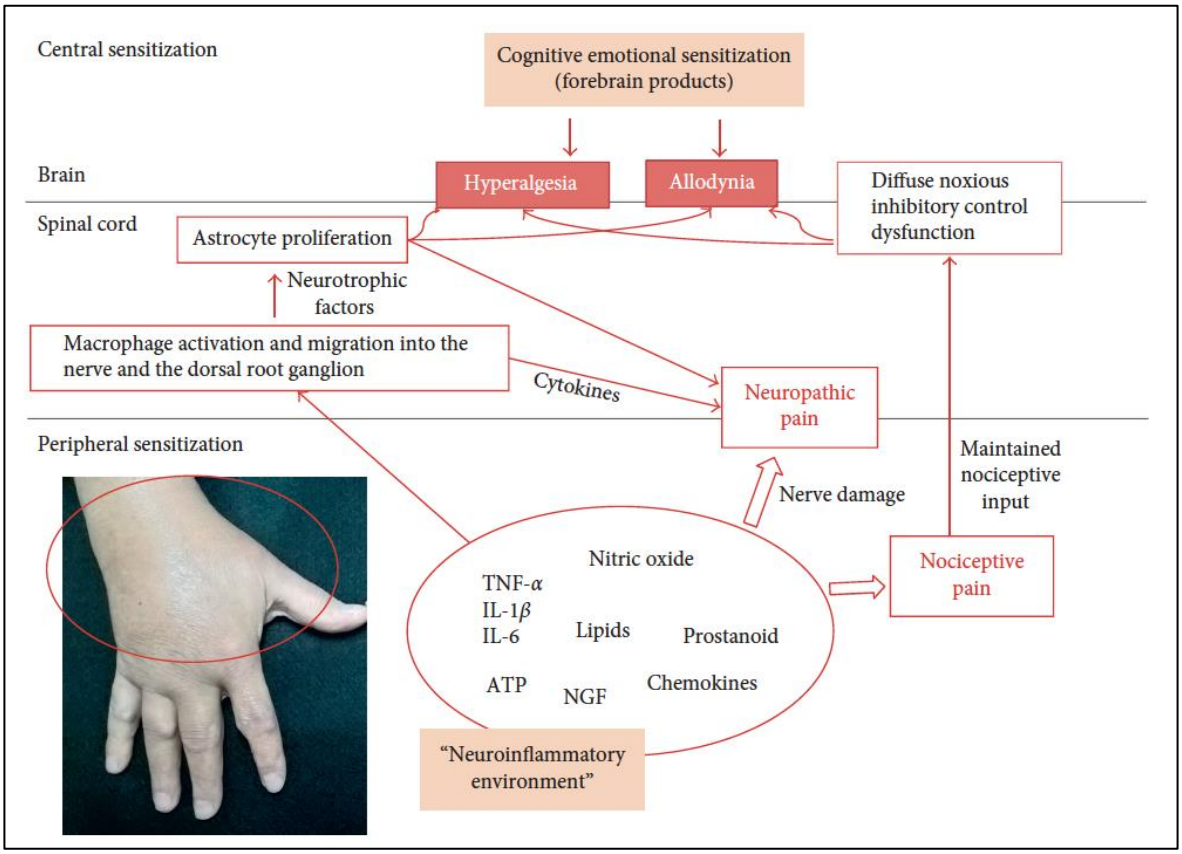


Figure 1.5 Pain mechanisms involved in inflammatory arthritis – peripheral and central pain sensitisations

Source: Adapted from Salaffi F, Giacobazzi G and Di Carlo M[76] (used with permission under Creative Commons Attribution License)

1.3.5 Persistent pain in arthritis exists despite suppressed inflammation

Advances in diagnosis and management of RA, including earlier diagnosis, treat-to-target (T2T) management, and paradigm shift of early disease-modifying anti-rheumatic drugs (DMARDs) use have revolutionised the rapid resolution of synovitis and other disease-

related systemic process. During treatment, pain relief is also expected when synovitis is under control, and yet, chronic pain in arthritis is complex and is far from just getting the disease itself under control.

In 1979, the International Association of Pain (IASP) Subcommittee on Taxonomy originally endorsed pain to be defined as ‘an unpleasant sensory and emotional experience associated with actual or potential tissue damage or described in terms of such damage’[77]. However, within this definition of pain, it is lacking the connotation that pain can exist in the absence of any body tissue injury. Four decades later, between 2018-2020, a revised IASP definition of pain was conceptualised as ‘an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage’[78]. This important change in defining pain is a crucial step in acknowledging the concept of central pain mechanisms and biopsychosocial circumstances within the individuals living with chronic pain.

Despite the absence of clinical or serological indicators of disease inflammation, persistent pain in RA is common and remains an ongoing treatment conundrum, implying the rationale behind alternative pain pathways, either physiologically or non-physiologically[79-81]. Central pain processing is augmented in both OA and RA and can occur distally from the original affected joints[82]. Heightened responsiveness of the somatosensory nerve fibres exists in surrounding peri-articular structures and other musculoskeletal structures[82-85]. In addition, response to pressure stimuli seems to dampen in RA, as evident by the heightened level of pressure pain detection in a widespread body distribution in response to cold stimuli in RA, but not in pain-free individuals[68, 86-88]. This impaired pressure pain threshold was shown to be further mediated by poor sleep, anxiety, and low mood[68, 87, 88].

Furthermore, neuropathic pain can co-exist in OA and RA, even in the absence of disease activity or activation of the classical nociceptive pain processing[89]. For instance, painDETECT questionnaire is widely used in musculoskeletal research studies examining

neuropathic and non-inflammatory pain[90-94]. In chronic, symptomatic knee OA, a correlation was found between higher scores from both the modified painDETECT questionnaire and positive quantitative sensory tests in nearly half of the study cohort, implying the presence of central pain sensitisation[93]. Similarly, in RA, there was a high proportion of study participants reporting high level of pain using the visual analogue score (VAS) and high painDETECT scores (R-squared statistics, R^2 of 0.757), implying the co-existence of neuropathic, non-inflammatory pain in a subset of patients with well-controlled disease[90].

As we now understand, central sensitisation exists in RA. This phenomenon is evident by the presence of mismatch between disease activity and pain symptoms, as well as the hyperexcitability of the central pain pathway in both articular and non-articular structures[85]. More importantly, this non-inflammatory type of pain is often comparable to the pain-related symptoms of undifferentiated chronic pain syndromes such as fibromyalgia or chronic fatigue syndrome[95-97]. In such non-inflammatory pain conditions, there is lack of nociceptive pain pathway and often, the central pain processing is augmented, as manifested by widespread hyperalgesia, allodynia, and dysregulation in pain inhibition[96, 98].

Studies have found that in RA, patients with centrally augmented chronic pain often fulfills the classification criteria of fibromyalgia. The prevalence of fibromyalgia and RA is relatively similar to that of the general population, but over time, as the RA disease progresses, the prevalence of fibromyalgia increases[99, 100]. These patients often have heightened pain sensitivity and more widespread, as well as adverse health outcomes such as poor sleep, impaired mental health, and poor quality of life[68, 101-106]. Frequently, fibromyalgia is independent of the standard measures of disease activity in RA, such as the 28-joint disease activity score (DAS28). In this circumstance, the higher DAS28 score is supported by higher scoring of the subjective components of DAS28, namely the tender joint count (TJC) and VAS of global health (VAS-GH)[68, 86, 96]. Recognising the mismatch between

disease activity and persistent pain has pawned detailed research into the contributing components of the disease activity scoring, objectively and subjectively. Studies in well-controlled, non-erosive RA cohorts have shown a large discrepancy between TJC and swollen joint count (SJC) and the corresponding ratio of TJC and SJC, resulting in abnormally high DAS28 scores in patients with persistent pain[101, 103, 107]. A study by McWilliams and colleague demonstrated that pain was predicted by high scoring of DAS28-P index, defined as "the proportion of DAS28 contributed by the patient-reported components"[108]. In this study, approximately three quarters of the study cohort reported partially improved or worsening pain at 12 months[108].

Overall, pain and inflammation in arthritis are mutually dependent, not only through the peripheral nociceptive pathway, but also through the central pain processing. It is pertinent for clinicians and researchers to understand how pain is measured in arthritis, especially over the natural history of disease and its progression. Next, pain assessment in arthritis is discussed, followed by a review of the current understanding of examining pain over time using various longitudinal analytical methods.

1.4 Pain Assessment in Arthritis

Patient-reported questionnaires on pain symptoms are commonly used by clinicians and researchers in the assessment of pain in arthritis. Either standalone or repeatedly over a period, these self-reported pain reports are crucial to capture the patient's perspective of their lived experience of chronic pain and its variability in time and space.

Some questionnaires cover other pain-related outcome measures, including pain-related symptoms, physical function, psychosocial status, quality of life and socioeconomic indicators, pertinent to any musculoskeletal diseases. In detail, there are simple one-dimensional pain scales that are easy and quick to use, such as VAS, numerical rating scale

(NRS) and verbal rating scale (VRS)[109-113]. These scales can be used for repeated measurements but may not be useful to capture more comprehensive picture of the pain experience. More complex, multidimensional pain questionnaires are available for use. In RA and OA, some of these questionnaires include McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), the Saint-Antoine Pain Questionnaire, the West Haven-Yale Multidimensional Pain Inventory (WHYMPI), Short Form-36 Bodily Pain Scale (SF-36 bodily pain scale), Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP), Arthritis Impact Measurement Scales (AIMS), the Western Ontario McMaster Questionnaire (WOMAC) and Rheumatoid Arthritis Pain Scale (RAPS)[59, 114-124]. Different aspects of pain measured in arthritis can be obtained depending on the desired outcome of interest, as summarised in Table 1.3.

These pain assessment tools are comprehensive, yet routine use in clinical care is not common, and unlikely to be feasible in a busy practice. Furthermore, if used in the usual clinical settings, responses to these questionnaires may be subject to recall bias and selective memory, as patients may be more likely to report any recent unpleasant pain experiences outside of the clinical setting. As such, in RA, when assessing disease activity, higher DAS28 may be subject to higher level of painful joints and overall poorer global health, even if the RA is stable. Overestimating the DAS28 score may lead to unnecessary treatment escalation and poor pain management in the longer term. The mismatch between disease activity and pain symptoms in this subgroup of patients with persistent pain highlights the importance of systematically capturing the pain experience over time, as well as capturing real-time pain impact. With this knowledge, we can then appropriately address their pain management and disease-related treatment decision and more importantly, intervene early when the window of opportunity arises throughout the disease course.

Pain scales		Health-related quality of life scales		Pain location scales	Site-specific scales
Unidimensional	**Multidimensional**	**Generic**	**Disease specific**		
Verbal rating scale (VRS) Visual Analogue Scale (VAS) Numerical Rating Scale (NRS) Faces Pain Rating Scale Thermometer Pain Scale (TPS)	McGill Pain Questionnaire (MPQ) Short-Form MPQ (SF-MPQ) Brief Pain Inventory (BPI) Chronic Pain Grade Questionnaire (CPGQ) West Haven-Yale Multidimensional Pain Inventory (WHYMPI) Rheumatoid Arthritis Pain Scale (RAPS)	36-Item Short-Form Health Survey (SF-36) European Quality of Life-5 Dimensions (EQ-5D) Sickness Impact Profile (SIP) Nottingham Health Profile (NHP)	Arthritis Impact Measurement Scales (AIMS) Arthritis Impact Measurement Scales (AIMS2)	Formal Joint Count Regional Pain Scale (RPS)	Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Hip Disability and Osteoarthritis Outcome Score (HOOS) Disabilities of the Arm, Shoulder, and Hand (DASH) Questionnaire

Table 1.3 Scales used for pain assessment in arthritis

Source: Adapted and restructured from Salaffi F, Giacobazzi G and Di Carlo M[76] (used with permission under Creative Commons Attribution License)

More often, unidimensional pain measurement tool is inadequate to measure the whole entity of pain experience in arthritis. In an ideal situation, continuous assessment of pain pattern in terms of its temporal and spatial distributions should occur, and more multidimensional pain assessment tools can potentially facilitate this type of longitudinal and comprehensive assessment. However, these multidimensional pain measurements are tedious and unappealing for repetitive use with less user engagement in data collection over time. Patient reported outcomes (PROs) are being adopted in clinical medicine and research given that PROs are important key measures of disease activity and treatment response derived from the patient's perspective on their disease outlook[125]. Such data collection attracts considerable interest in the longitudinal data capture of pain symptoms

and pain-related PROs, and hence, the review of the rationale and advances of real-time granular data capture is discussed next.

1.4.1 The history and advances of real-time daily data capture on pain experience

Diary use in capturing symptoms and co-occurring events over time for familial periodic paralysis was first described in the early 1920s[126]. This method of data collection was not fully taken up until half a century later, when the "Experience Sampling Method (ESM)" was first developed by Csikszentmihalyi and Larson[127]. The ESM originates from Csikszentmihalyi's work on flow theory, which describes the relationship of an individual's state of consciousness with performance and activity engagement spatially and temporally[128]. The ESM has a primary emphasis on recording momentary information of an individual's physical and mental state experience in natural settings, by completing questionnaire handouts in writing, and at various pre-arranged time points[127]. In order to capture such data, signalling devices were used to provide prompting to the users to record any subjective phenomena or events relevant to the researchers' outcome of interest[127]. Such research based on the ESM approach, the signalling devices used ranged from pagers, computerised pocket calculators, computerised wrist watch, tape recorder, to customised-built portable Epson computer, which could emit more than 100 signals over just a week period[127].

With the implementation of this first era of wireless telecommunication devices and computers, the ESM was considered an advancement in real-time longitudinal data capture compared to the traditional fixed time-point questionnaire assessment. Subsequently, the ESM led the paradigm shift in capturing real-time self-reported health data, particularly in studies of psychology and physical illness[129-131]. Besides capturing momentary data on people's daily activity and symptoms, methods of collecting data on behavioural and physiological measures were already achievable since the mid 1990s, with the use of

ambulatory monitoring devices (e.g., blood pressure and heart rate monitor) and physical activity sensors[132, 133].

Major shift in the paradigm of real-time data capture was subsequently seen with the development of "Ecological Momentary Assessment (EMA)" by Stone and Schiffman[134]. The main components of the EMA are defined by "real-time collection of data about momentary states, collected in the natural environment, with multiple repeated assessments over time"[135]. The EMA is a framework aggregating all the above historical methods described earlier, and ultimately, aims to seek and to integrate all the common features of real-time data collection from a methodological perspective. One would certainly ask why we need such an onerous framework, given that we have good reliability and validity of the traditional self-report tools used in many health studies.

1.4.2 The risks of recall bias and selective memory in pain assessment

The concept of 'peak and end rule' as reviewed in detail by Fredrickson is one of the most important determining factors influencing recall bias, which is a common challenge when using the traditional self-reporting questionnaires[136]. This concept largely explains the theory of cognitive heuristics. In any co-occurring events, most human beings are prone to subjectively conclude an emotional experience based on a few selected past moments which are usually at its most intense period (and within this concept, it is known as 'the peak') and the eventual outcome (also known as 'the end')[136]. A study by Redelmeier and Kahneman further confirmed the concept of this cognitive heuristic theory[137]. In this study, included patients underwent minimally invasive procedures such as colonoscopy and lithotripsy and their overall pain experience was strongly correlated with recall of the most intense pain during the procedure and towards the end (last few minutes) of the procedure[137]. Another study by Redelmeier and Kahneman confirmed that the duration of the procedure did not significantly affect their overall pain perception unless the final

impression of the procedure was correlated with unpleasantness[138]. It suggests that human being's memory is naturally selective and is susceptible to bias. Conceptually, regardless of the duration, if there were cumulative moments of pain experience or other negative affect, these would conflict with the temporal monotonicity and the overall perception of its aversiveness[139].

Although the studies as outlined above confirmed the concept of the 'peak and end rule', it is important to note that the study results were derived within a controlled research environment over a very short period. These study results may not be generalisable to real-time capture of pain symptoms as seen in ethnographic studies. For instance, an important study by Stone and his colleague is worth to discuss, with retrospective self-reported daily pain experience captured over a week period in patients with RA using the EMA framework[140]. In this study, in a natural environment with minimal interruption to daily activities, study participants recorded their daily mood, arthralgia and impact on function using a 7-point scale rating in a booklet[140]. In addition to pain symptoms, the status quo activities, whereabouts, and co-occurring events relevant to disease impact were recorded[140]. Regular prompts for data entry were given throughout the day using pre-programmed watch worn by the study participants[140]. Concurrently, during this study period, one rheumatology visits for each participant occurred with the intention to record participant's recall of their average joint pain and other pain symptoms[140]. Study results derived from these 32 dedicated participants confirmed the validity of the combined 'peak-end effect' of pain experience being more superior than the recalled averaging of pain alone[140].

For patients and clinicians, such momentary methods of data collection can be unappealing and intrusive. However, if the data collection can be made easier, and with very minimal interference with daily living, EMA applicability in momentary snapshot of certain PROs in studies of chronic health conditions can yield the most natural, temporally rich longitudinal

data. Next, the discussion on how the current emerging technologies can assist with this method of real-time data capture is presented.

1.4.3 The case for mobile health application in capturing real-time patient-generated health data

The rapid growth of mobile phone use has led to a parallel rise in mHealth studies since the late 90s. Broadly, mHealth is known as the 'emerging mobile communications and network technologies for health care'[141]. Almost 63% of the global population were mobile phone users in 2016, and the number of users continued to rise in 2019 and beyond[142]. Mobile phone is no longer considered as a simple tool for communication. There is ongoing development and advancement of internal capabilities within a smartphone, such as webpage access, video and voice recording, geo-positioning system (GPS), and a user-friendly touch screen interface. These built-in technologies have revolutionised the adoption of its convenience in participatory health care in the wider society, irrespective of age, geography, and socio-economic status[143].

Currently, mHealth has ongoing influence in advancing health care research and practice, with wider adoptions seen in health-care decision-making using artificial intelligence (AI) system, in remote monitoring and care of those with chronic medical conditions, and in acute medical care[144]. Seamless built-in sensors, efficiently powered devices, and wireless networks found within the mHealth devices have enhanced the feasibility of using these devices in health care and research[145]. The notion of empowering patients to take charge of their own health is encouraged in this setting, with strong focus in place for self-management and self-monitoring of disease activity and impact in many chronic health conditions[145]. Overall, mHealth has great potentials and benefits not only for individuals and health care professionals, but also for the wider society and key stakeholders in health care and research.

One in two smartphone users have at least one downloaded mHealth application (app), according to a cross-sectional survey study conducted on 1604 smartphone users in the United States[146]. According to this study, although health apps were considered by most users to be beneficial in self-care awareness and health improvement, nearly half of these smartphone users discontinued using the apps, primarily due to data entry burden, gradual lack of interest, and costs[146]. Despite the large influx of downloaded health-related mobile apps, user engagement remains a challenge. On a similar note, some of these readily available health-related mobile apps are not rigorously assessed and validated, resulting in lack of its usefulness in clinical setting. A systematic review on RA-related mobile apps has shown that most of these were commercially available apps that are downloadable from the iTunes and Google Play App Stores[147]. Most of these apps were found to be less feasible for longer term use in self-care and self-monitoring of disease activity and impact in RA[147].

From the health care research perspective, there is a large uptake of mHealth application in most medical and social science studies to gather longitudinal biopsychosocial-focused PROs[135, 148]. For instance, such high level of data granularity can be analysed using multilevel modelling, which has the advantages of assessing the between-individual and within-individual processes within the outcome of interest[149]. Next, key points in mHealth implementation and feasibility in rheumatology research is discussed.

1.4.4 The case for longitudinal pain symptom tracking using smartphones

Part of the expanding horizon in digital health is the development of patient-generated health data using mHealth and globally, the number of mHealth-related studies in chronic health conditions is increasing.

For instance, a feasibility study of EMA-based data collection of chronic pain in individuals with FM using electronic diaries embedded within smartphones has shown the success in maintaining user engagement, even in those with less digital literacy[150]. Furthermore, a

smartwatch study from the United Kingdom (UK), known as the Knee Osteoarthritis: Linking Activity and Pain (KOALAP) study, has demonstrated groundwork success in consistent study user participation and engagement in providing real-time data on symptoms, alongside passive data collection on physical activity[151]. Concurrent active and passive data collection from individuals in real-time is possible in mHealth studies, providing consistent user engagement.

In longitudinal tracking of symptoms experienced by individuals with RMDs, in terms of pain experience, invaluable information in great details on symptom variability over time can be obtained[152]. For instance, observational studies on low back pain (LBP) have largely identified distinct subgroups of varying pain trajectories using clustering analysis, allowing identification of at-risk individuals with varying pain level and better tailoring of pain management. Pain experience fluctuates, and the dynamic of pain states can either be static or volatile. For example, a study by Axen and colleagues identified distinct subgroups of pain trajectories in patients with LBP, described as either sustained or fluctuating rather than a single form of pain trajectory for the whole study population[153]. Apart from understanding pain intensity, quality and distribution, the temporal entity, frequency, and the transition between pain states are all relevant to capture the overall lived experience of pain in real-time, and this is a positive direction for better understanding of chronic pain in arthritis.

The term ‘e-Rheumatology’, as introduced by El Miedany, has clearly defined the benefits of mHealth technologies in enhancing the continuum of care between the health care providers and patients with RMDs[154]. In a study of smartphone users with early RA, greater level of self-improvement was seen in the users, with the development of positive outlook and personal confidence in self-management of symptoms and medication[155]. Apart from this, a 3-month feasibility study of using smartphone devices for self-assessment of RA has shown good correlation between patients’ perception of their disease activity and clinicians’ objective outcome measures in the clinic[156]. More importantly, these findings suggest that

by using mHealth technologies, self-management for those with arthritis is possible, further complementing clinical care with the systematic input from the patients and objective disease markers. Patient engagement and involvement in their health care, alongside with clinicians' assessment and support, will undoubtedly provide a clearer snapshot of the disease impact in their lives and their treatment outcome. Advancing consumer technology such as smartphones and wearables has enabled patients to easily track their daily symptoms and its variability in trend through time[157]. These patient-generated data can be beneficial if integrated in clinical practice, particularly in remote monitoring of daily symptoms reported by the patient outside of the clinic rooms[157].

1.5 Key Measures of Pain-Related Health Outcomes in Arthritis

Health status is often perceived as a state of physical health, mental health, and social well-being. The natural history of an individual's health status is an ever-changing metric in one's life course, ranging from a state of wellness or being asymptomatic to illness onset and its trajectory, if present[158-160]. In arthritis, persistent pain may have an impact on medication use and the overall health status. For instance, in RA, pain is regarded as the highest priority in health domains for improvement and in measures of treatment efficacy[161-164]. Additionally, higher levels of pain experienced in RA were strongly correlated with poor quality of life, as measured by decline in the overall health perception, increased dependency and need for support, and decline in biopsychosocial functioning[165-170].

Overarchingly, looking into the intertwined relationship between the complex dynamics of pain and arthritis and the negative corollary health outcomes that followed allows us to look deeper into the overall impression of the well-being of the person living with arthritis, especially in those with persistent pain. To date, we are yet to have standardised methods to examine pain longitudinally. In addition, little is known of the temporal relationship

between pain and health status of patients living with arthritis, and more importantly, of how trajectories of pain-related health outcomes translate into attributable burden on treatment, and ultimately, the downward effects on morbidity and mortality of those at-risk individuals with persistent pain.

1.6 Thesis Aims

The research groundwork in this thesis aims to explore the current gaps and understanding of longitudinal characteristics of pain in arthritis. Such exploration of longitudinal data on pain-related PROs captured in large observational studies, including mHealth studies, may unravel new insights for patients, clinicians, and key stakeholders of the phenomenology of persistent pain and its attributable burden on treatment and adverse health outcomes in arthritis. Part of the research work in this thesis has a dedicated focus on RA.

The research questions in this thesis include the following:

- What are the methods used to longitudinally assess pain in arthritis in the current literature?
- What are the benefits and challenges in using digitalised health-related data in clinical care and research? In the context of harnessing patient-generated health data obtained from smartphones, what are the key outcomes that can be identified when assessing pain trajectory and variability using intensive longitudinal methods?
- In RA, by using the validated pain score, disease activity composite score and other relevant PROs, can we identify individuals at risk of developing persistent pain trajectory? What is the impact of persistent pain on medication use and important health outcomes in RA, such as hospitalisation events, mortality risk and causes of death (COD)?

To address the abovementioned research questions, this thesis aims:

- To explore the current literature on the methods and research tools used in longitudinal characteristics of pain in arthritis
- To synthesise the benefits and challenges of digitalised health-related data in providing evidence-based care
- To explore daily pain variability in inflammatory and non-inflammatory RMDs using intensive longitudinal methods
- To determine the longitudinal effects of pain and health outcomes and the attributable burden of pain-related health status on treatment and adverse health outcomes in RA

1.7 Thesis Outline

This thesis is presented as a 'combination thesis' format with most thesis chapters written in publication format. Throughout my thesis, as a guide for the reading, each section starts with an overview of the chapter(s) and ends with a summary of the section. In parallel, each chapter also starts with a preface and a summary of the chapter. Chapters with published articles and manuscript submitted for publication in peer-reviewed journals are presented. This thesis comprises three sections as described below.

Section 1 includes a chapter on my systematic scoping review (Chapter 2) on the current literature in longitudinal assessment of pain in RA. A manuscript for this review has been prepared and submitted to a peer-reviewed journal.

Section 2 consists of two chapters:

- A narrative review (Chapter 3) on the benefits and challenges of harnessing digitalised health-related data in providing real-world evidence. This review has been published in a peer-reviewed journal.
- Chapter 4 focuses on the work completed on an mHealth study, *Cloudy with a Chance of Pain*. I joined the research team in this study in the early stage of my candidature in Manchester, UK. This is a UK nationwide smartphone study investigating the association between pain and weather. I joined the study when the primary analysis of interest had just begun, following the conception and design of the study, the study app, pilot study and the completion of recruitment and data collection. As part of the research team for the primary analysis, I was partly involved in the data cleaning of the baseline questionnaires. I also contributed as a co-author for the first three publications of this study during my candidature (included in the Appendices). Specific to my thesis, I focused on using a data subset from this study of study participants with inflammatory and non-inflammatory RMDs. Using this data subset, I analysed day-to-day pain trajectory and variability using intensive

longitudinal methods. This exploratory analysis is presented as a traditional thesis chapter. The statistical code scripts are included in this thesis in the Appendices.

Section 3 consists of three chapters:

- Chapter 5 focuses on the work completed on an early RA cohort with the aim to examine the components of the DAS28-P index as a discriminatory measure of treatment response in this study cohort. This work has been published in a peer-reviewed journal. The dataset is described in detail in this chapter.
- Chapter 6 focuses on the work completed on the dataset of study participants with RA derived from the Australian Rheumatology Association Database (ARAD). The aim was to examine the longitudinal effects of pain on sociodemographic indicators and medication use. A manuscript for this analysis has been published in a peer-reviewed journal. The dataset is described in detail in this chapter.
- Chapter 7 focuses on the work continued from Chapter 6 using the dataset of study participants with RA derived from the ARAD. The aim was to examine the longitudinal effects of pain on hospitalisations, mortality risk and causes of death. A manuscript for this analysis has been prepared and submitted to a peer-reviewed journal.

Next, Chapter 8 focuses on the overall discussion of the research groundwork in this thesis. In this chapter, the key findings, significance, and implications of the findings in clinical care and research are discussed. In addition, the strengths and limitations of the research are discussed. Research skills attained during my candidature and the impact of Covid-19 on the progress of my candidature are discussed.

To conclude this thesis, as presented in Chapter 9, the overall conclusion and future directions are discussed.

Section 1: Literature Review

Overview of Section 1

Section 1 of this thesis consists of one chapter, Chapter 2, which forms the primary driver for the study design and conception behind most of the research work conducted in this thesis.

Chapter 2 presents the systematic scoping review of the current literatures in examining the longitudinal characteristics of pain in RA, an autoimmune RMD mostly featured in my research work in pain trajectories. This chapter also presents the research gap identified through this scoping review, therefore, forming the basis of my research work in RA in this thesis.

Chapter 2: Longitudinal Characteristics of Pain in Arthritis: A Scoping Review

2.1 Preface

The systematic scoping review presented in this chapter addresses the first part of the research questions in this thesis. This review aims at evaluating the current literatures on methods used to longitudinally assess pain in RA, an autoimmune RMD mostly featured in my research work on pain trajectories presented in Chapters 5, 6, and 7 in this thesis.

This scoping review is presented in a manuscript format, which has been submitted to a peer-reviewed journal, the BMC Rheumatology. The Statement of Authorship is included. The numbered references in this manuscript have been reformatted accordingly in this thesis, corresponding to the Reference thesis chapter. To end, I present the chapter summary and research gaps identified from this scoping review.

2.2 Statement of Authorship

Statement of Authorship

Title of paper	Longitudinal Characteristics of Pain in Rheumatoid Arthritis: A Scoping Review
Publication status	Submitted for Publication
Publication Details	Pisaniello, H.L., Nairne-Nagy, J., Lester, S., Whittle, S.L., Hill, C.L. Longitudinal Characteristics of Pain in Rheumatoid Arthritis: A Scoping Review. BMC Rheumatology Impact Factor: 2.451

Principal Author

Name of principal author	Huai Leng Pisaniello
Contribution to the Paper	Contributed to the study design and conception, the development of literature search strategy, database search, data acquisition and management. Contributed to the title and abstract screening, review of full-text articles, data extraction and quality assessment of included studies for the review. Contributed to the data analysis and interpretation of study results. Contributed to the preparation, drafting and revision of the manuscript for publication.
Overall percentage (%)	75%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 8 April 2023

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. The candidate's stated contribution to the publication is accurate (as detailed above);
ii. Permission is granted for the candidate in include the publication in the thesis; and
iii. The sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Jaiden Nairne-Nagy
Contribution to the Paper	Contributed to the title and abstract screening, review of full-text articles, data extraction and quality assessment of included studies for the review. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft.
Signature	Date: 28/11/2023

*For Jaiden Nairne-Nagy

Name of Co-Author	Susan Lester
Contribution to the Paper	Contributed to the supervision of the literature search and review process. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 18/04/2023

Name of Co-Author	Samuel L. Whittle
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 18/04/2023

Name of Co-Author	Catherine L. Hill
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 18/04/2023

2.3 Manuscript: Longitudinal Characteristics of Pain in Rheumatoid Arthritis: A Scoping Review

Submitted to *BMC Rheumatology*

Title:

Longitudinal Characteristics of Pain in Rheumatoid Arthritis: A Scoping Review

Authors:

Huai Leng Pisaniello, MBBS, FRACP[1,2],

Jaiden Nairne-Nagy[1],

Susan Lester, BSc[2,3],

Samuel L. Whittle, MBBS (Hons), MClinEpi, FRACP[2,3],

Catherine L. Hill MBBS MD MSc FRACP[1,2,3]

Affiliations:

[1]Discipline of Medicine, Faculty of Health and Medical Sciences, The University of Adelaide, South Australia, Australia.

[2]Rheumatology Research Group, Basil Hetzel Institute for Translational Health Research, The Queen Elizabeth Hospital, Woodville South, South Australia, Australia.

[3]Rheumatology Unit, The Queen Elizabeth Hospital, Woodville South, South Australia, Australia.

Corresponding Author:

Huai Leng Pisaniello

Rheumatology Research Group, Basil Hetzel Institute for Translational Health Research, The Queen Elizabeth Hospital, Woodville South, South Australia, Australia

Email: huaileng.pisaniello@adelaide.edu.au

ORCiD No: 0000-0002-0425-1697

ABSTRACT

Objective: This scoping review aimed to identify the methods applied in longitudinal analysis of pain in rheumatoid arthritis (RA) and to explore the relationship between pain trajectory and potential predictors.

Methods: Three electronic databases (Medline, Embase and PsycInfo) were systematically searched (study inception to May 2022). Eligibility criteria included adults aged ≥18 years, with RA and with longitudinal methodology and analysis designed to examine repeated measures of pain (≥3 timepoints). Relevant studies identified in the search were independently screened by two reviewers and extracted data from the included studies were critically synthesised.

Results: Of the 1,400 unique records retrieved, 22 studies met the inclusion criteria. All except one were prospective cohort studies, using a wide variety of pain scales and frequency of pain measurements. Nearly half of the included studies described changes in pain over time as group-level mean pain. Two studies were able to stratify pain trajectories into distinct pain subgroups using clustering analysis and another two studies applied novel dynamic modelling approaches in exploring pain variability in RA. Female gender, ethnic minorities, smoking history, and low education level were associated with higher pain levels, with female gender and younger age being significantly correlated with frequent pain fluctuations over time.

Conclusion: Stable pain trajectories in RA with minimal fluctuations of pain level were observed, regardless of the heterogeneity of the study designs and outcomes. Future research incorporating standardised pain measurements (pain scales and frequency) and

novel pain trajectory and pain variability methodologies is required in examining pain experience in RA.

Keywords: rheumatoid arthritis, pain, pain trajectory, pain variability

Word count: Abstract: 250**;** Main text: 4,405

INTRODUCTION

Rheumatoid arthritis (RA), widely known as a systemic autoimmune-mediated inflammatory disease, results in joint inflammation and pain and can be associated with other extra-articular manifestations[10]. Disease remission in RA with better long-term outcomes is now achievable through early diagnosis and a 'treat-to-target' strategy, alongside the booming era of biologic disease-modifying anti-rheumatic drugs (DMARDs) use in managing advanced RA. Despite effective disease control using DMARDs and with minimal radiological progression, a large proportion of patients with RA report ongoing pain[80, 108]. In a meta-analysis of long-term data of patient-reported outcomes in patients with early RA, pain, fatigue, and functional disability did not significantly improve over time, even up to 5 years of established RA on treatment[171]. Persistent pain in RA is complex and has multiple contributors including noxious inputs associated with disease inflammation and joint damage, as well as altered central pain processing changes, or commonly known as the nociplastic phenomenon. Persistent pain in RA can cause negative impact on quality of life, physical function and psychological health[172].

Pain experience in RA is highly variable, depending on the disease course and severity, treatment response, comorbid pain conditions as well as the demographics and psychosocial factors within the individuals[108, 172, 173]. Traditionally, research into the longitudinal characteristics of pain summarises change in pain over time as the population-averaged pain level, and interestingly, in RA, reports of mean pain scores have remained similar despite the improved use of DMARDs over the past two decades[174]. Unfortunately, these mean pain scores provide little to minimal meaningful information when it comes to

differentiating patients with different pain experience over time or assistance in tailoring individual pain management and RA treatment.

In psychology research, in the field of personality assessment, summarising group-level patterns of the outcome of interest across individuals is traditionally applied, also commonly known as the nomothetic approach. In contrast, experience sampling methods which are common research tools used in the modern idiographic approach, consist of real-time capture of the outcome of interest over time for each individual, allowing a within-individual longitudinal analysis of variability[175]. In detail, the idiographic approach allows the examination of individual differences over time in the temporal and behavioural distribution and the trajectory process of the outcome of interest[175]. Similarly, when it comes to describing pain experience, the alternative to the nomothetic approach in averaging pain levels over time across individuals is to apply an idiographic approach of examining within-individual patterns of pain longitudinally. Furthermore, subsequent stratification of these individual-level pain measures longitudinally into distinct subgroups of pain over time can be identified using appropriate trajectory analytical techniques. For instance, work in identifying subgroups or phenotypes in osteoarthritis (OA) has expanded over the past decade, which coincides with one of the top research priorities by the European Alliance of Associations for Rheumatology (EULAR)[176]. A recent systematic review by Wieczorek and colleagues identified 44 published studies which reported trajectory analysis of clinical outcomes in non-surgical knee and hip OA[177]. Of interest, in this review, stable disease trajectory in knee and hip OA was seen, despite the heterogeneity in the study outcomes, number of distinct subgroups and the phenotype of the individuals within each subgroup[177].

For longitudinal research in pain experience, in addition to identifying pain trajectories, examining the fluctuations of pain intensity or pain variability within the pain trajectory is equally important. Capturing the ebb and flow of the pain experience within the individual will allow visualisation of the dynamic process of chronic musculoskeletal pain in greater

detail, as seen in some non-inflammatory rheumatic conditions such as OA and fibromyalgia[178, 179].

The nature of the research in examining and characterising pain in RA longitudinally is unclear, and to our knowledge, no reviews summarising the methodological aspects and results of such studies have been published. The aims of this scoping review are to examine the methods used in longitudinal analysis of pain in RA; to analyse the relationship between pain trajectory and potential predictors; and to identify gaps in the current literature in this area in RA.

MATERIALS AND METHODS

Methodologic framework

The methods designed and used in this scoping review followed the guidelines proposed by Arksey and O'Malley as well as by Levac and colleagues[180, 181]. This review was written in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR)[182].

Search strategy

Medline, Embase and PsycInfo databases were systematically searched for relevant studies using keywords linked to three main concepts: rheumatoid arthritis, pain and keywords pertaining to longitudinal pain. The search was conducted by the primary author (HLP), in consultation with an academic librarian. Published articles were searched from inception and up to May 27, 2022. The final search strategy for each database is available in the Supplementary Files.

Eligibility criteria

Published articles of adult populations (aged 18 years and older) with RA and with pain data with at least 3 timepoints or more of repeated measures of pain and were analysed longitudinally were included in the review. In addition to published articles, our review also included studies in abstract or conference proceeding format. Only studies published in English were included. Pertinent to our study aims of interest in longitudinal methods in analysing pain experience over time in RA, we excluded studies examining treatment response in RA (including studies examining the placebo/nocebo effect) and studies on pain due to non-RA causes such as malignancy, neurological, pregnancy or menstruation, abdominal or urological, vascular, and post-surgical pain. In addition, studies published in the form of editorials, opinions, review articles, and lab and basic science studies were excluded.

Study selection

The articles were retrieved, and duplicates were removed using EndNote v20 software (Clarivate analytics). Using the Covidence platform (a systematic review software, Veritas Health Innovation, Melbourne, Australia; available at www.covidence.org), articles were independently screened by 2 reviewers (HLP and JNN) first by title and abstract, followed by full text. Disagreements were resolved by consensus discussion or a third reviewer (SL). HLP manually searched reference lists of relevant articles.

Data extraction

A data extraction table was piloted and amended accordingly to the aims of the scoping review. Data extraction was performed by the primary author (HLP). Data charted were then verified within the research team against the original articles for accuracy and for any discrepancies identified, these were resolved via consensus discussion.

The data extracted included the primary author, year of study, country, study design, sample size and the sample population. With regards to the longitudinal pain analysis, extracted

data included the type of pain scale, length of follow-up, number of measurements over time, and the information on the statistical analysis and software used. Baseline characteristics of the included study cohorts were extracted, which included the duration of disease, early or established RA, gender, age, and treatment. Significant pain-related covariates/confounders or predictors (including sociodemographic factors, clinical factors, patient-reported measures, or psychological factors) associated with the primary longitudinal analysis of the pain were also extracted. A formal risk of bias assessment was not done in keeping with standard methodology for a scoping review[182].

Statistical Analysis

Study characteristics of the included studies were summarised using descriptive statistics, with the aim to broadly synthesise the nature of current research on this review topic. First, type of pain scale, frequency of measurements over time and study duration were described separately and compared between the included studies. Second, the study cohorts for each study were categorised into early or established RA subgroups, providing an overview of the demographics of the study participants and treatment exposure. Third, methods used to analyse pain trajectory in each study were summarised based on the analytical methods used in describing the overall pain trajectory and variability over time. The pain trajectory outcomes were summarised based on the reported descriptive analysis and stratification of pain trajectory subgroups where applicable. Covariates and confounders identified within the pain trajectory outcomes and predictors of pain were summarised. Data extraction, aggregation and charting was performed in Microsoft Excel and tabulated summaries of the review findings were presented.

RESULTS

The literature search using the three electronic databases identified a total of 1,399 studies. One additional study was identified through the manual search of the reference lists of

included studies. After the removal of 360 duplicates, 1,040 studies were screened by the title and abstract. From this initial screen, 995 studies were excluded; the remaining 45 studies were screened by full-text. A total of 22 studies were eligible for the final review (Figure 1).

Study characteristics

Of the 22 included studies, 21 were prospective cohort studies[183-203], and one study was a retrospective analysis of registry data (Table 1 and Table 2)[204]. A comprehensive summary for the data extracted for these included studies is provided in the Supplementary File.

From these 22 included studies, there were 19 published articles[183, 184, 186-190, 192-195, 197-204] and 3 abstracts[185, 191, 196]. It is important to note that some studies used similar study cohorts with either articles being published by different authors or at different times. For instance, the authors, Affleck and Ward, used the same study cohort in their articles published at different times[183, 184, 201, 202]. Additionally, the use of similar study cohort was noted in some studies, such as those studies by Harries and McWilliams (both studies used ERAN, Early Rheumatoid Arthritis Network)[191, 193], and by Lotsch and Sandberg (both studies used EIRA, a Swedish-based Epidemiological Investigation of Rheumatoid Arthritis)[196, 204].

All included studies used self-reported pain questionnaires, with wide variation in types of pain scales. The study publication years ranged from 1991 to 2022. The follow-up duration in most of the studies, when specified, ranged from 7 days to 10 years. Most of the included studies had regular timepoints of repeated measures of pain, ranging from daily (n = 7, with one study having 7 repeated measures daily)[183, 184, 186, 189, 190, 197, 198], weekly (n = 3)[185, 187, 199], fortnightly (n = 1)[201], monthly (n = 1)[200] and at least 6 monthly (n = 4, with 1 studies having 6 monthly pain measures in the first year, and annually thereafter)[192, 202, 203]. Five other studies had irregular timepoints of pain measures[188, 194-196, 204]. Two studies, although the frequency of repeated pain measures was not specified, it was assumed that there were

at least 3 timepoints or more of the measures of pain based on their study durations[191, 193]. Six studies included patients with early RA[188, 191-193, 196, 204], 11 studies included patients with established RA[183-185, 187, 193-195, 200-203] and 6 studies included patients with RA of unspecified disease duration[186, 189, 190, 197-199]. In these studies, at least 6 different instruments were used to assess pain, with visual analogue score (VAS; presented in many forms including 0-100mm scale, 0-10 scale, descriptive scale of 'no pain' to 'very severe pain') and numerical rating scale (NRS; range 0-10) being used most frequently.

Methods and statistical models

Nearly half of the included studies involved an overall descriptive summary of group-level mean pain to describe change in pain over time[183, 184, 188, 190-194, 203]. Four studies used repeated measures analysis of variance (ANOVA), a method used to compare the means across an outcome variable based on the repeated measures of pain[186, 188, 194, 195]. Multilevel models were used in 2 studies, which is useful to identify the population-average trajectory and the random variation around the average trajectory using one set of parameters for the study sample[197, 199]. These models consider the variability from the individual (random effects), alongside the standard population-level mean. Time-lag and autocorrelation models were used in 2 studies, allowing the examination of preceding pain scores and successive pain score differences[183, 200]. In terms of assessing the patterns of fluctuations of pain over time, 2 studies identified components of pain variability using Markov regime switching dynamic regression model and non-linear damped oscillator model[189, 198]. We identified 2 studies which performed clustering analysis and captured distinct subgroups of pain trajectories, using growth mixture modelling and machine learning unsupervised analysis by fitting a Gaussian mixture model to the Pareto density estimation, a kernel density estimator[193, 204].

Pain trajectories

Across all included studies, the overall summary of pain experience in RA is that of initial pain variability at the start of the study, but over time, the trajectory of pain remained stable with minimal fluctuations of pain level, regardless of the pain intensity. However, results between studies also highlight the heterogeneity of the study cohort and the timepoints of repeated measures of pain taken for the study duration.

Stratifying pain trajectories into distinct subgroups of pain is possible, using clustering analysis methods that are commonly used in the OA literature[177]. For instance, in the paper by McWilliams and colleagues, the clustering analysis was performed on 3 different cohorts: the early RA cohort using the Early RA Network (ERAN), the biologics and non-biologics cohorts in established RA using the British Rheumatology Biologics Register for biologics (BSRBR-biologics) and non-biologics (BSRBR-non-biologics) respectively[193]. Discrete persistent pain (59-79%) and resolving pain (19-27%) subgroups were captured in each of these cohorts, but in the early RA cohort, there was an additional third subgroup of persistently low pain (23%)[193]. In the paper by Lotsch and colleagues, 3 distinct pain subgroups were captured: low-, medium- and high-persistent pain intensity[204].

Covariates/confounders and predictors of pain

Overall, the covariates and predictors studied within the pain trajectories in these included studies were heterogenous. Age and gender were the two most used variables in these included studies. In addition, the other categories included were sociodemographic and socioeconomic variables (ethnicity, education, smoking history), mental health (mood, depression), physical function and disability, and biological factors such as composite objective measures of disease activity (inflammatory markers).

In detail, among the sociodemographic factors assessed in some studies, female gender and younger age were associated with more frequent pain fluctuations[187]. In the study by Kumaradev and colleagues, when assessing the transition from early RA to established disease, the differences in pain evolution mainly derived as a function of age, gender and

ethnicity, such that younger patients, males and Caucasians demonstrated lower pain in the later phase of disease[192]. Higher education was associated with lower pain level in the early disease phase, with no changes throughout the disease course[192]. Smoking history was an identified risk factor for persistent pain in early and established RA study cohorts[193]. Similarly, in the study by Wolfe and colleagues, higher pain levels were seen in women, ethnic minorities, smokers, and those with low education status[203].

Some of the included studies also identified an association with other covariates including anxiety, depression, mood, fatigue, and pain coping strategies. A study by Sandberg and colleagues identified a correlation between pain and disease activity and inflammation in the first 3 months, with loss of the correlation throughout the remainder of the study period, with tender joint counts being significantly more correlated with pain than swollen joint counts(SJCs)[196]. Interestingly, beyond 3 months after RA diagnosis, there were sustained higher correlations between pain, TJCs and patient global assessment (PGA) compared to inflammatory markers throughout the study period, highlighting the importance of evaluating other causes of pain, besides the disease inflammation[196].

DISCUSSION

Our scoping review included 22 studies that investigated longitudinal characterisation of the identified pain trajectories of RA in adult populations. Our review highlights significant heterogeneity between the studies. Overall, descriptive summary statistics were predominantly used to describe the change in pain over time. Two studies were able to identify homogenous pain subgroups within their corresponding study cohorts. Furthermore, two other studies examined the pain variability using novel dynamic process modelling approaches, allowing better characterisation of pain components within the pain variability over time. The covariates and predictors included in the studies were highly heterogenous, and these include several sociodemographic and psychosocial factors.

Nearly all the included studies were prospective population-based cohorts, with varying study duration ranging from 1 week to 10 years. Similarly, the pain scales used in each included study, although all were self-reported, differed across studies, alongside using varying frequency of repeated measures of pain during the study period. This review highlights female preponderance in the study cohorts for all included studies. Future research should consider a standardised approach of assessing pain in clinical trials and epidemiologic studies.

In some studies which identified distinct subgroups of pain trajectories, different pain intensity subgroups were identified, alongside the persistence of the pain level throughout the trajectories. These findings also highlight the importance of examining covariates or predictors within the homogenous subgroups of pain rather than relying on the average trend for the whole study population.

A consistent finding in this review was initial variability in pain scores among people with RA, followed by an overall stable pain trajectory in RA, regardless of the disease course and treatment, suggestive of minimal fluctuations of pain over time. Similarly, this finding is consistent with other observational studies examining osteoarthritic pain in the hip and knee joints, with a large proportion of study participants remained in a 'state of inertia' in terms of their symptom and radiographic progression over time[205-207]. Furthermore, in this review, the index pain measurement used in the trajectory analyses would be based on the first timepoint of study enrolment or assessment rather than the actual time of disease onset, which is a challenge when interpreting the trajectory results. Biases from missing data and attrition due to patient-related fatigue in data contribution or loss in follow-up may influence the true estimate of the pain trajectories, especially in those who discontinued the study due to persistently high pain level[208]. Determining the likelihood and rate of pain progression often remains an ongoing clinical dilemma for many patients and health care providers, as symptomatic RA can be considered as a chronic inflammatory disease process. To provide

an accurate assessment of pain trajectory in symptomatic RA, index pain measured from the disease onset, where possible, and high levels of study engagement and retention may be necessary.

In studies with more frequent timepoints of pain measurements taken, pain variability was described better in higher resolution, with different pain components within the variability being described. For instance, in the study by Schneider and colleagues examining the temporal dynamics of pain, 4 different components of pain variability were captured using a novel Markov regime switching regression model[198]. These include mean pain, mean amplitude of pain transition, mean persistence of pain states and dominance of pain states, allowing precise characterization of pain variability and examination of the potential association of different pain components with important pain-related predictors[198]. In addition, the study by Finan and colleagues identified pain prediction being more accurate over time by using the non-linear damped oscillatory model[189]. It is noteworthy to consider other factors that could explain the similarity of pain levels between distant time periods in these studies. Most frequently used pain scales in these studies, such as the VAS and NRS, are considered as ordinal scales of measure. Ordinal scale assumes similar difference between the pain levels in these scales in order, but the interval or magnitude of difference between each pain level may not be constant, such as the temperature range[209]. The application of appropriate statistical methods when using such scales remains contentious, especially in analysing data of repeated measurements between individuals[209]. Pain perception is subjective to individuals, and therefore, the magnitude of pain level experienced may differ, even if the pain level reported between individuals is the same. When conducting research on evaluating pain experience, recall bias remains an issue, as the pain-related questionnaires often assume the individuals to accurately recall their pain or to summarise their pain level over a pre-defined period[136, 210]. Furthermore, cross-cultural differences in pain reporting exists, therefore, an important factor to consider when asking the individuals to express their pain level using culturally and linguistically appropriate pain assessment instruments[211].

Despite the advancement in the treatment for RA, some patients continue to experience persistent pain, despite minimal disease activity and progression. Pain in RA is multidimensional, and in those with mismatch between minimal disease inflammation and ongoing pain, the contribution of either the central nociplastic pain processing or other pain pathway such as neuropathic pain is noteworthy[7, 68]. For example, in a study by Ahmed and colleagues, pain was assessed using VAS and painDETECT, an instrument that is designed to measure non-inflammatory and centrally augmented pain, including neuropathic pain[90]. In this study, there was significant correlation between patients with neuropathic pain measured by painDETECT and increased total pain reported using VAS[90]. Although our scoping review excluded studies of neuropathic pain alone, one study by Sandberg and colleagues identified the importance of reviewing other causes of pain apart from joint inflammation, based on loss of correlation of pain and inflammatory markers after 3 months but persistently higher correlations between pain, TJCs and PGA[196]. In this scoping review, although not explicitly reported in the included studies, female predominance was prominent, partly highlighting the well-established natural history of RA, a chronic autoimmune musculoskeletal disease, that affects mostly women with a female to male ratio of 3:1[212, 213]. In parallel, gender disparity has been shown to influence the reporting of patient-reported outcome measures and pain experience when evaluating treatment efficacy, and as such, for trials of evaluating pain and treatment response in RA, female predominance is an important confounder to consider in the interpretation of the study results[214, 215].

Our scoping review aimed to highlight the existing work that has examined the longitudinal characteristics of pain in RA. Work on stratifying pain experience in RA into distinct subgroup is lacking in comparison to the volume of trajectory analyses in knee and hip OA[177]. Furthermore, our review identified novel dynamic process modelling approaches in examining pain variability in greater details. Future studies combining both the assessment of pain trajectory and the examination of pain variability within the identified trajectory may

provide better characterisation of the overall pain experience, short-term and long-term, for patients with RA, regardless of treatment and disease course.

In addition, our scoping review has shown a paradigm shift in longitudinal methods in characterizing pain experience in RA. Many studies in the early years used simple summary statistics in describing changes in pain over time, and in recent studies, novel and sophisticated statistical methods were applied to better characterise the pain trajectories and pain variability in RA. More importantly, advances in capturing pain variability are regarded as equally important in summarizing momentary pain experience in patients with chronic pain in general[216]. Better accuracy in describing pain experience alongside the disease course may facilitate more personalised treatment use and pain management for patients with different pain patterns[217, 218]. The composite measures of pain variability may provide insights into changes in pain over time and treatment response in different pain subgroups, potentially leading to more meaningful patient-centered endpoints in clinical trials[219-221]. For future research, apart from the core predictors identified from this scoping review, such as age, gender, and sociodemographic indicators, it is pertinent to carefully consider the type of pain scale and the corresponding statistical methods used. The use of the International System of Units in natural sciences has shown the benefits of standardised measurements and in contrast, the progress of harmonising outcome measures in health sciences remains an ongoing issue, as seen in the heterogeneity in patient-reported outcome measures (PROMs) used in assessing pain[222]. In a recent systematic review by Georgopoulos and colleague has demonstrated group-level similarity in the scoring of pain threshold criterion, known as the Patient Acceptable Symptom State (PASS), when different PROMS for knee pain were standardised[222, 223]. Such harmonisation of different PROMS in measuring pain outcome may provide further insights when comparing pain trajectory studies. Additionally, this may provide an alternative analytical method when analysing pain trajectory using different PROMs within the same study dataset, where possible. Apart from data harmonisation, in the same systematic review, the use of PASS as the proxy of pain threshold criterion has shown that pain acceptability at an individual level is largely

dependent on the patient demographics, disease progression, baseline pain score, and treatment characteristics[222, 223]. Other predictors that are relevant to any longitudinal pain assessment in RA may include the effect of composite disease activity score, important co-occurring pain symptoms such as fatigue and mood and the overall wellbeing of the individual.

This scoping review has some limitations. We did not perform a risk of bias assessment for the included studies, and therefore, we could not comment on the quality of the studies, in particular the reporting of the methods used and study results. Second, we were not able to quantitatively summarise the findings of the included studies, due to the heterogeneity of the statistical methods used and the instruments used in the study for pain assessment. Timepoint frequency of pain measures differed significantly between studies, resulting in challenges to summarise the longitudinal characteristics of pain in RA. In addition, our scoping review did not include studies which examined non-inflammatory persistent pain in those with disease remission in RA, and therefore, our findings in this review cannot be extrapolated to this specific population. We did not include articles published in languages other than English, or unpublished data, a potential bias in the study inclusion and exclusion process. Although this scoping review is limited to the longitudinal characterisation of pain in RA, it is equally important to consider the outcome measure of pain derived from the placebo effect in determining treatment response in RA. A systematic review by Abdullah on 165 randomised controlled trials (RCTs) of treatment response in RA has demonstrated a significant trend towards pain reduction in the placebo arm compared to treatment arm in RA (overall placebo effect size, defined as standardised mean difference from baseline, of 0.28 [95% confidence interval (CI) 0.19, 0.37])[224]. Such findings highlight the importance of considering the placebo phenomenon when evaluating treatment efficacy in RA, especially in RCTs with therapeutic ceiling effects[224, 225]. Overall, future studies with coherent study population, statistical methods and pain scales may be necessary for accurate comparison between studies describing pain experience in RA.

In conclusion, our scoping review highlighted the heterogeneity of describing pain experience in RA across studies. Overall, pain trajectories appear to be stable over time in RA, with initial fluctuations of pain seen at the beginning of the study period. Significant heterogeneity in describing the relationships between different covariates and predictors in terms of sociodemographic, disability and psychosocial factors was observed. Examining the pain variability within the course of pain over time is possible using novel dynamic process modelling approaches, suggesting the potential for future work to incorporate both analyses of pain trajectories and pain variability within the identified trajectories.

Acknowledgement

We would like to thank Ms Anna Holasek, the reference and academic librarian of South Australia Health Library Service for the Queen Elizabeth Hospital and Basil Hetzel Institute for her help and advice with the literature search for this scoping review.

Author Contributions

All authors were involved in the drafting process and contributed intellectual contents for this article. All authors approved the final version to be submitted for publication. Dr. Pisaniello had full access to the study data and materials and takes responsibility for the data integrity and the accuracy of the interpretation of the data analysis.

Study conception and design: Pisaniello, Lester, Whittle, Hill.

Acquisition of data: Pisaniello, Nairne-Nagy, Lester.

Analysis and interpretation of data: Pisaniello, Nairne-Nagy, Lester, Whittle, Hill.

Conflict of Interest

The authors declared having no conflict of interest in relation to this manuscript.

Funding Information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

This manuscript formed part of HLP's Ph.D. work, for which HLP was the recipient of the Ken Muirden Overseas Training Fellowship from the Arthritis Australia, an educational grant funded by the Australian Rheumatology Association, and HLP also received the Australia Postgraduate Scholarship from the University of Adelaide.

References

1. Sparks JA. Rheumatoid Arthritis. Ann Intern Med. 2019 Jan 1;170(1):ITC1-ITC16. DOI: 10.7326/AITC201901010.
2. Lee YC, Cui J, Lu B, Frits ML, Iannaccone CK, Shadick NA, et al. Pain persists in DAS28 rheumatoid arthritis remission but not in ACR/EULAR remission: a longitudinal observational study. Arthritis Res Ther. 2011 Jun 8;13(3):R83. DOI: 10.1186/ar3353.
3. McWilliams DF, Zhang W, Mansell JS, Kiely PD, Young A, Walsh DA. Predictors of change in bodily pain in early rheumatoid arthritis: an inception cohort study. Arthritis Care Res (Hoboken). 2012 Oct;64(10):1505-13. DOI: 10.1002/acr.21723.
4. Carpenter L, Barnett R, Mahendran P, Nikiphorou E, Gwinnutt J, Verstappen S, et al. Secular changes in functional disability, pain, fatigue and mental well-being in early rheumatoid arthritis. A longitudinal meta-analysis. Semin Arthritis Rheum. 2020 Apr;50(2):209-19. DOI: 10.1016/j.semarthrit.2019.08.006.
5. Walsh DA, McWilliams DF. Pain in rheumatoid arthritis. Curr Pain Headache Rep. 2012 Dec;16(6):509-17. DOI: 10.1007/s11916-012-0303-x.
6. Forslind K, Hafstrom I, Ahlmen M, Svensson B, Group BS. Sex: a major predictor of remission in early rheumatoid arthritis? Ann Rheum Dis. 2007 Jan;66(1):46-52. DOI: 10.1136/ard.2006.056937.
7. Welsing PM, Fransen J, van Riel PL. Is the disease course of rheumatoid arthritis becoming milder? Time trends since 1985 in an inception cohort of early rheumatoid arthritis. Arthritis Rheum. 2005 Sep;52(9):2616-24. DOI: 10.1002/art.21259.
8. Conner TS, Tennen H, Fleeson W, Barrett LF. Experience Sampling Methods: A Modern Idiographic Approach to Personality Research. Soc Personal Psychol Compass. 2009 May 1;3(3):292-313. DOI: 10.1111/j.1751-9004.2009.00170.x.
9. Conaghan PG, Kloppenburg M, Schett G, Bijlsma JW, committee Eoah. Osteoarthritis research priorities: a report from a EULAR ad hoc expert committee. Ann Rheum Dis. 2014 Aug;73(8):1442-5. DOI: 10.1136/annrheumdis-2013-204660.
10. Wieczorek M, Rotonda C, Guillemin F, Rat AC. What Have We Learned From Trajectory Analysis of Clinical Outcomes in Knee and Hip Osteoarthritis Before Surgery? Arthritis Care Res (Hoboken). 2020 Dec;72(12):1693-702. DOI: 10.1002/acr.24069.
11. Allen KD. The value of measuring variability in osteoarthritis pain. J Rheumatol. 2007 Nov;34(11):2132-3.
12. 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. DOI: 10.1002/art.21407.
13. Arksey H OML. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8:19-32.
14. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci. 2010 Sep 20;5:69. DOI: 10.1186/1748-5908-5-69.
15. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018 Oct 2;169(7):467-73. DOI: 10.7326/M18-0850.
16. Affleck G, Tennen H, Urrows S, Higgins P. Individual differences in the day-to-day experience of chronic pain: a prospective daily study of rheumatoid arthritis patients. Health Psychol. 1991;10(6):419-26. DOI: 10.1037//0278-6133.10.6.419.
17. Affleck G, Urrows S, Tennen H, Higgins P. Daily coping with pain from rheumatoid arthritis: patterns and correlates. Pain. 1992 Nov;51(2):221-9. DOI: 10.1016/0304-3959(92)90263-B.

18. Bailly F JC, Servy H, Fautrel B, Gossec L. Variability of pain levels is explained by self-reported disease acitvity in rheumatoid arthritis and axial spondyloarthritis: A 3-months prospective study of 165 patients. Annals of Rheumatic Diseases 2018;77:538.
19. Cruise CE, Broderick J, Porter L, Kaell A, Stone AA. Reactive effects of diary self-assessment in chronic pain patients. Pain. 1996 Oct;67(2-3):253-8. DOI: 10.1016/0304-3959(96)03125-9.
20. Drouet J, Gossec L, Jacquemin C, Fautrel B, Foltz V, Gandjbakhch F, et al. Fluctuation of pain is frequent in rheumatoid arthritis and axial spondyloarthritis: A 12 weeks prospective study of 165 patients. Joint Bone Spine. 2022 May;89(3):105306. DOI: 10.1016/j.jbspin.2021.105306.
21. Eberhard A, Bergman S, Mandl T, Olofsson T, Rydholm M, Jacobsson L, et al. Predictors of unacceptable pain with and without low inflammation over 5 years in early rheumatoid arthritis-an inception cohort study. Arthritis Res Ther. 2021 Jun 14;23(1):169. DOI: 10.1186/s13075-021-02550-7.
22. Finan PH HE, Amazeen PG, Butner J, Zautra AJ, Tennen H. Oscillations in daily pain prediction accuracy. Nonlinear dynamics, psychology and life sicences. 2010;14(1):27-46.
23. 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. DOI: 10.1093/rap/rkac021.
24. Harries L MD, Young A, Kiely PDW, Walsh DA. Prediction of future pain by DAS28-P in patients with early rheumatoid arthritis: The ERAN cohort. Rheumatology (UK). 2014;1:i71-i2.
25. Kumaradev S, Roux C, Sellam J, Perrot S, Pham T, Dugravot A, et al. Socio-demographic determinants in the evolution of pain in inflammatory rheumatic diseases: results from ESPOIR and DESIR cohorts. Rheumatology (Oxford). 2022 Apr 11;61(4):1496-509. DOI: 10.1093/rheumatology/keab562.
26. McWilliams DF, Dawson O, Young A, Kiely PDW, Ferguson E, Walsh DA. Discrete Trajectories of Resolving and Persistent Pain in People With Rheumatoid Arthritis Despite Undergoing Treatment for Inflammation: Results From Three UK Cohorts. J Pain. 2019 Jun;20(6):716-27. DOI: 10.1016/j.jpain.2019.01.001.
27. Odegard S, Finset A, Mowinckel P, Kvien TK, Uhlig T. Pain and psychological health status over a 10-year period in patients with recent onset rheumatoid arthritis. Ann Rheum Dis. 2007 Sep;66(9):1195-201. DOI: 10.1136/ard.2006.064287.
28. Roche PA, Klestov AC, Heim HM. Description of stable pain in rheumatoid arthritis: a 6 year study. J Rheumatol. 2003 Aug;30(8):1733-8.
29. Sandberg M SS, Altawil R, Alfredsson L, Klareskog L, Lampa J. The relationship between pain and inflammation in early rheumatoid arthritis: Long-term follow-up in the Swedish population-based EIRA and SRQ register. Scandinavian Journal of Rheumatology. 2014;127:4-5.
30. 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):813-22. DOI: 10.1016/j.pain.2012.01.001.
31. Schneider S, Junghaenel DU, Ono M, Stone AA. Temporal dynamics of pain: an application of regime-switching models to ecological momentary assessments in patients with rheumatic diseases. Pain. 2018 Jul;159(7):1346-58. DOI: 10.1097/j.pain.0000000000001215.

32. Smith BW, Zautra AJ. The effects of anxiety and depression on weekly pain in women with arthritis. Pain. 2008 Aug 31;138(2):354-61. DOI: 10.1016/j.pain.2008.01.008.
33. van Dartel SA, Repping-Wuts JW, van Hoogmoed D, Bleijenberg G, van Riel PL, Fransen J. Association between fatigue and pain in rheumatoid arthritis: does pain precede fatigue or does fatigue precede pain? Arthritis Care Res (Hoboken). 2013 Jun;65(6):862-9. DOI: 10.1002/acr.21932.
34. Ward MM. Are patient self-report measures of arthritis activity confounded by mood? A longitudinal study of patients with rheumatoid arthritis. J Rheumatol. 1994 Jun;21(6):1046-50.
35. Ward MM, Leigh JP. The relative importance of pain and functional disability to patients with rheumatoid arthritis. J Rheumatol. 1993 Sep;20(9):1494-9.
36. Wolfe F, Michaud K. Assessment of pain in rheumatoid arthritis: minimal clinically significant difference, predictors, and the effect of anti-tumor necrosis factor therapy. J Rheumatol. 2007 Aug;34(8):1674-83.
37. Lotsch J, Alfredsson L, Lampa J. Machine-learning-based knowledge discovery in rheumatoid arthritis-related registry data to identify predictors of persistent pain. Pain. 2020 Jan;161(1):114-26. DOI: 10.1097/j.pain.0000000000001693.
38. Bastick AN, Wesseling J, Damen J, Verkleij SP, Emans PJ, Bindels PJ, et al. 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):e32-9. DOI: 10.3399/bjgp15X688129.
39. Felson D, Niu J, Sack B, Aliabadi P, McCullough C, Nevitt MC. Progression of osteoarthritis as a state of inertia. Ann Rheum Dis. 2013 Jun;72(6):924-9. DOI: 10.1136/annrheumdis-2012-201575.
40. Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SM, Boers M, Cats HA, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9. DOI: 10.1136/ard.2008.096164.
41. Losina E, Collins JE. Forecasting the future pain in hip OA: can we rely on pain trajectories? Osteoarthritis Cartilage. 2016 May;24(5):765-7. DOI: 10.1016/j.joca.2016.01.989.
42. Kim TK. Practical statistics in pain research. Korean J Pain. 2017 Oct;30(4):243-9. DOI: 10.3344/kjp.2017.30.4.243.
43. B.L. F. Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. Cognition Emotion. 2000;14(4):577-606.
44. Haefeli M, Elfering A. Pain assessment. Eur Spine J. 2006 Jan;15 Suppl 1(Suppl 1):S17-24. DOI: 10.1007/s00586-005-1044-x.
45. Peacock S, Patel S. Cultural Influences on Pain. Rev Pain. 2008 Mar;1(2):6-9. DOI: 10.1177/204946370800100203.
46. McWilliams DF, Walsh DA. Pain mechanisms in rheumatoid arthritis. Clin Exp Rheumatol. 2017 Sep-Oct;35 Suppl 107(5):94-101.
47. Zhang A, Lee YC. Mechanisms for Joint Pain in Rheumatoid Arthritis (RA): from Cytokines to Central Sensitization. Curr Osteoporos Rep. 2018 Oct;16(5):603-10. DOI: 10.1007/s11914-018-0473-5.
48. Ahmed S, Magan T, Vargas M, Harrison A, Sofat N. Use of the painDETECT tool in rheumatoid arthritis suggests neuropathic and sensitization components in pain reporting. J Pain Res. 2014;7:579-88. DOI: 10.2147/JPR.S69011.
49. van Vollenhoven RF. Sex differences in rheumatoid arthritis: more than meets the eye. BMC Med. 2009 Mar 30;7:12. DOI: 10.1186/1741-7015-7-12.

50. Wolfe AM, Kellgren JH, Masi AT. The epidemiology of rheumatoid arthritis: a review. II. Incidence and diagnostic criteria. Bull Rheum Dis. 1968 Nov;19(3):524-9.
51. Maranini B, Bortoluzzi A, Silvagni E, Govoni M. Focus on Sex and Gender: What We Need to Know in the Management of Rheumatoid Arthritis. J Pers Med. 2022 Mar 20;12(3). DOI: 10.3390/jpm12030499.
52. Sokka T, Toloza S, Cutolo M, Kautiainen H, Makinen H, Gogus F, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. Arthritis Res Ther. 2009;11(1):R7. DOI: 10.1186/ar2591.
53. Stone AA, Obbarius A, Junghaenel DU, Wen CKF, Schneider S. High-resolution, field approaches for assessing pain: Ecological Momentary Assessment. Pain. 2021 Jan;162(1):4-9. DOI: 10.1097/j.pain.0000000000002049. Cited in: Medline.
54. Martini CH, Yassen A, Krebs-Brown A, Passier P, Stoker M, Olofsen E, et al. A novel approach to identify responder subgroups and predictors of response to low- and high-dose capsaicin patches in postherpetic neuralgia. Eur J Pain. 2013 Nov;17(10):1491-501. DOI: 10.1002/j.1532-2149.2013.00329.x. Cited in: Medline.
55. Treister R, Honigman L, Lawal OD, Lanier RK, Katz NP. A deeper look at pain variability and its relationship with the placebo response: results from a randomized, double-blind, placebo-controlled clinical trial of naproxen in osteoarthritis of the knee. Pain. 2019 Jul;160(7):1522-8. DOI: 10.1097/j.pain.0000000000001538. Cited in: Medline.
56. Mun CJ, Suk HW, Davis MC, Karoly P, Finan P, Tennen H, et al. Investigating intraindividual pain variability: methods, applications, issues, and directions. Pain. 2019 Nov;160(11):2415-29. DOI: 10.1097/j.pain.0000000000001626. Cited in: Medline.
57. Schneider S, Junghaenel DU, Ono M, Broderick JE, Stone AA. III. Detecting Treatment Effects in Clinical Trials With Different Indices of Pain Intensity Derived From Ecological Momentary Assessment. J Pain. 2021 Apr;22(4):386-99. DOI: 10.1016/j.jpain.2020.10.003. Cited in: Medline.
58. Stone AA, Broderick JE, Schneider S, Schwartz JE. Expanding options for developing outcome measures from momentary assessment data. Psychosom Med. 2012 May;74(4):387-97. DOI: 10.1097/PSY.0b013e3182571faa. Cited in: Medline.
59. Dekker J. Harmonization of pain scores - interesting and relevant findings. Osteoarthritis Cartilage. 2023 Jan;31(1):9-10. DOI: 10.1016/j.joca.2022.08.017.
60. Georgopoulos V, Smith S, McWilliams DF, Steultjens MPM, Williams A, Price A, et al. Harmonising knee pain patient-reported outcomes: a systematic literature review and meta-analysis of Patient Acceptable Symptom State (PASS) and individual participant data (IPD). Osteoarthritis Cartilage. 2023 Jan;31(1):83-95. DOI: 10.1016/j.joca.2022.08.011.
61. Abdullah N. Placebo effect in the treatment of rheumatoid arthritis: a systematic review and meta-analysis of randomized controlled trials. Paper presented at: EULAR ARD; 2015.
62. Bechman K, Yates M, Norton S, Cope AP, Galloway JB. Placebo Response in Rheumatoid Arthritis Clinical Trials. J Rheumatol. 2020 Jan;47(1):28-34. DOI: 10.3899/jrheum.190008.

FIGURE AND TABLES LEGENDS

Figure 1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart demonstrating the systematic search strategy and screening process to capture included studies for the scoping review

Table 1: Description of the included studies in the scoping review

Table 2: Description of the pain instruments and frequency of measurements, methodologies, results (including covariates and predictors) in the included studies

Table 3: Key findings for the included studies

Figure 1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart demonstrating the systematic search strategy and screening process to capture included studies for the scoping review

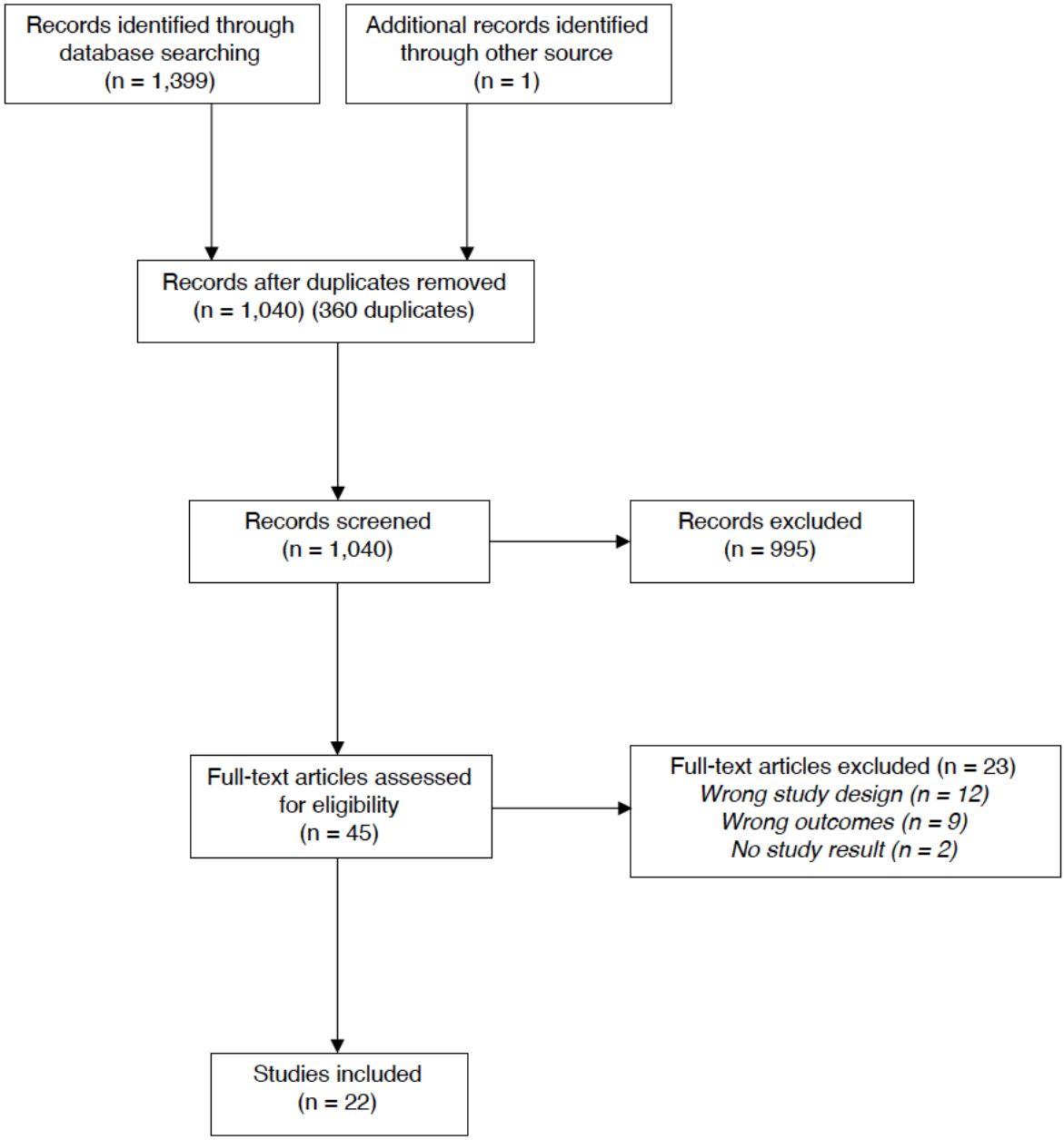


Table 1: Description of the included studies in the scoping review

Reference	Country	Study design	Sample size	Female gender, %	Mean age, years (SD)	Mean disease duration, years (SD)	RA treatment
Affleck et al, 1991[183]	US	Prospective cohort	47	76.60%	52.77 (12.79)	9.06 (7.66)	Not specified
Affleck et al, 1992[184]	US	Prospective cohort	75	71%	52.7 (12.5)	8.9 (7.7)	Not specified
Bailly et al, 2018 (abstract)[185]	France	Prospective cohort	86*	81%	48.7 (12.7)	8 (8.8)	bDMARDs
Cruise et al, 1996[186]	US	Prospective cohort	18*	71%**	52.4 (12.3)	Not specified	Not specified
Drouet et al, 2022 (letter to the Editor)[187]	UK	Prospective cohort	86*	63%	Median 45 years	Median 8 years	bDMARDs
Eberhard et al, 2021[188]	Sweden	Prospective cohort	232	70.30%	60.5 (14.6)	Median 7 months (IQR 5-10)	cDMARDs and bDMARDs
Finan et al, 2010[189]	US	Prospective cohort	170	74%	55.2 (13.3)	Not specified	Not specified
Gandrup et al, 2022[190]	UK	Prospective cohort	20	70%	Median 58.5 (IQR 48-64)	Not specified	Not specified
Harries et al, 2014 (abstract)[191]	UK	Prospective cohort	1,236	Not specified	Not specified	Not specified	Not specified
Kumaradev et al, 2022[192]	France	Prospective cohort	810	76%	48.1 (12.2)	0.6 (0.7)	NSAIDs; cDMARDs and bDMARDs
Lotsch et al, 2020[204]	Sweden	Retrospective cohort	288	73%	52.2 (12.3)	Not specified	Not specified
McWilliams et al, 2019[193]	UK	Prospective cohort	9,493	66% (ERAN), 77% (BSRBR - biologics), 75% (BSRBR - non-biologics)	57 (ERAN), 57 (BSRBR - biologics), 61 (BSRBR - non-biologics)	0.8 [1.3] (ERAN), 13 [10] (BSRBR - biologics), 10 [11] (BSRBR - non-biologics)	cDMARDs and bDMARDs
Odegard et al, 2007[194]	Norway	Prospective cohort	149	76%	50.2 (12.5)	2.2 (1.2)	NSAIDs; glucocorticoids; cDMARDs and bDMARDs
Roche et al, 2003[195]	Australia	Prospective cohort	120	72%	58.75 (12.58)	15.27 (10.09)	Not specified
Sandberg et al, 2014 (abstract)[196]	Sweden	Prospective cohort	1802	Not specified	Not specified	Not specified	Not specified
Schneider et al, 2012[197]	US	Prospective cohort	30*	86%	56 (11)	Not specified	Not specified

Schneider et al, 2018[198]	US	Prospective cohort	24*	84%	57 (SD 13.12)	Not specified	Not specified
Smith et al, 2008[199]	US	Prospective cohort	82*	Not specified	62.1 (SD 7.3)	Not specified	Not specified
Van Dartel et al, 2013[200]	The Netherlands	Prospective cohort	228	64%	56.7 (SD 10.6)	Median 10 (IQR 6-17)	Glucocorticoids; cDMARDs and bDMARDs
Ward 1994[201]	US	Prospective cohort	24	92%	Median 46 (IQR 28-73)	Median 3 (IQR 0.2-20)	Not specified
Ward 1993[202]	US	Prospective cohort	305	83%	53.9 (SD13.8)	14.1 (SD 9.0)	Not specified
Wolfe et al, 2007[203]	US	Prospective cohort	12090	76%	60.09 (SD 13.61)	14.27 (SD 11.12)	Not specified

All included studies were published journal articles unless specified.
**Subset of the total study sample size, **Aggregated result*
bDMARDs: biologic DMARDs, cDMARDs: conventional DMARDs, DMARDs: disease modifying anti-rheumatic drugs, ERAN: Early Rheumatoid Arthritis Network, BSRBR: British Society for Rheumatology Biologic Registry, IQR: interquartile range, NSAIDs: non-steroidal anti-inflammatory drugs, RA: rheumatoid arthritis, UK: United Kingdom, US: United States

Table 2: Description of the pain instruments and frequency of measurements, methodologies, results (including covariates and predictors) in the included studies

Reference	Early/established RA	No. of measurements over time	Study duration	Pain scale	Longitudinal characteristics of pain	Covariates and predictors analysed
Affleck et al, 1991[183]	Established RA	Daily	75 days	VAS (no pain - very severe pain)	Mean pain over time; pain trajectory	Disease-related; mental health; physical function/disability
Affleck et al, 1992[184]	Established RA	Daily	75 days	VAS (no pain - very severe pain)	Mean pain over time; pain trajectory	Gender; mental health; physical function/disability
Bailly et al, 2018 (abstract)[185]	Established RA	Weekly	3 months	NRS (0-10)	Mean pain over time; pain variability	Disease-related
Cruise et al, 1996[186]	Not specified	7 times daily	7 days	7-point scale (0-6 with 0 = not at all, 3 = moderate, and 6 = extremely)	Mean pain over time	Mental health; physical function/disability
Drouet et al, 2022 (letter to the Editor)[187]	Established RA	Weekly	12 weeks	NRS (0-10)	Mean pain over time; pain trajectory	Age; disease-related; gender
Eberhard et al, 2021[188]	Early RA	Irregular	5 years	VAS (0-100mm)	Mean pain over time; pain trajectory	Age; disease-related; gender; physical function/disability; quality of life
Finan et al, 2010[189]	Not specified	Daily	29 days	NRS (101-point)	Pain variability (non-linear damped oscillator model)	Mental health
Gandrup et al, 2022[190]	Not specified	Daily	85 days	NRS (0-10)	Mean pain over time	Disease-related
Harries et al, 2014 (abstract)[191]	Early RA	Not specified*	3 years	SF-36 bodily pain	Mean pain over time; pain trajectory	Disease-related; mental health; physical function/disability; quality of life
Kumaradev et al, 2022[192]	Early RA	Irregular	10 years	SF-36 bodily pain and VAS (0-100)	Mean pain over time	Age; ethnicity; education; gender
Lotsch et al, 2020[204]	Early RA	Irregular	5 years	VAS (0-100mm)	Mean pain over time; pain trajectory with clustering analysis (unsupervised machine learning)	Disease-related
McWilliams et al, 2019[193]	Early RA and established RA	Not specified*	3 years	SF-36 bodily pain	Mean pain over time; pain trajectory with clustering analysis	Disease-related; physical function/disability; smoking

					(growth mixture modelling)	
Odegard et al, 2007[194]	Established RA	Irregular	10 years	VAS (0-100mm)	Mean pain over time; pain trajectory	Gender; disease-related; mental health; physical function/disability
Roche et al, 2003[195]	Established RA	Irregular	77 months	VAS (0-10)	Mean pain over time; pain trajectory	Mental health
Sandberg et al, 2014 (abstract)[196]	Early RA	Irregular	5 years	VAS (0-100mm)	Pain trajectory	Disease-related
Schneider et al, 2012[197]	Not specified	Daily	28 days	VAS (101-point)	Pain variability	Mental health
Schneider et al, 2018[198]	Not specified	Daily	3 months	VAS (0-10)	Mean pain over time; pain variability (Markov regime switching dynamic regression modelling)	Mental health; physical function/disability
Smith et al, 2008[199]	Not specified	Weekly	11 weeks	Questionnaire 0-100 of average pain	Mean pain over time; pain trajectory	Mental health
Van Dartel et al, 2013[200]	Established RA	Monthly	12 months	NRS (0-10)	Mean pain over time; pain trajectory	Gender; mental health; physical function/disability
Ward 1994[201]	Established RA	Fortnightly	60 weeks	VAS (15cm) with 0 = no pain and 3 = very severe pain	Mean pain over time; pain trajectory	Disease-related; mental health; physical function/disability
Ward 1993[202]	Established RA	Biannually	9.5 years	VAS (15cm) with 0 = no pain and 100 = very severe pain	Mean pain over time; pain trajectory	Ethnicity; gender; disease-related; physical function/disability
Wolfe et al, 2007[203]	Established RA	Biannually	Not specified	VAS (0-10)	Mean pain over time; pain trajectory	Age; disease-related; ethnicity; education; gender; smoking status

NRS: numerical rating scale; RA: rheumatoid arthritis; SF-36: Short Form 36; VAS: visual analogue scale

Table 3: Key findings for the included studies

Reference	Key findings
Affleck et al, 1991[183]	Pain scores were highly correlated across successive days; more intense pain trajectory was associated with more active disease, depression, and disability.
Affleck et al, 1992[184]	Higher levels of pain were associated with less positive mood; improving pain trajectory was associated with overall better pain coping efforts.
Bailly et al, 2018 (abstract)[185]	Low fluctuations of pain were observed throughout the study period; pain fluctuations were highly correlated with self-reported flares and PGA of disease activity.
Cruise et al, 1996[186]	Mean pain remained stable over time; fatigue was positively correlated with pain; positive reactive affects were negatively correlated with pain intensity.
Drouet et al, 2022 (letter to the Editor)[187]	Although mean pain remained stable over time, pain fluctuations over time were frequent, with strong association with female gender and higher baseline PGA.
Eberhard et al, 2021[188]	Pain improvement was seen in the first 6 months but remained unchanged for the remainder of the study period; one third of patients had unacceptable pain at 5 years after inclusion, predicted by lower SJCs, higher VAS and higher PGA of disease activity at baseline; unacceptable pain with low inflammation was negatively associated with anti-CCP status.
Finan et al, 2010[189]	Pain prediction accuracy oscillated over time, with larger amplitude of oscillation at the start, indicating damping towards more accurate prediction; low negative affect and higher pain control were associated with better pain prediction accuracy.
Gandrup et al, 2022[190]	Mean pain and pain-related symptoms and variability were associated with higher odds of flare.
Harries et al, 2014 (abstract)[191]	Pain improvement was seen in the first year but remained unchanged for the remainder of the study period; DAS28-P significantly predicted the next year's pain.
Kumaradev et al, 2022[192]	In the transition from early to established RA, pain trajectory emerged as a function of age and ethnicity; younger age, being male and Caucasians were associated with lower pain in the latter disease phase; high education was associated with lower pain in the early disease phase, with no changes throughout the disease course.
Lotsch et al, 2020[204]	3 distinct pain trajectory subgroups were identified: low-, medium- and high-persistent pain intensity.
McWilliams et al, 2019[193]	Mean pain improved over time in each cohort but remained >1 standard deviation worse than the UK general population throughout the study period; 2 pain trajectory subgroups were identified in each cohort: discrete persistent pain and resolving pain; 1 additional pain subgroup was identified in the early RA cohort - persistently low pain; higher disability and smoking history were risk factors for persistent pain trajectories in each cohort.
Odegard et al, 2007[194]	The pain level was explained longitudinally by anxiety, disease activity, physical function, and female gender. Anxiety, not depression, was significantly associated with the course of pain.
Roche et al, 2003[195]	Mean pain remained stable over time, with no overall change in pain sensation, affect and emotional quality.
Sandberg et al, 2014 (abstract)[196]	Pain in RA was correlated with disease inflammation at diagnosis, but the correlation with objective disease markers disappeared after 3 months; TJCs were significantly correlated with pain over time than SJCs.
Schneider et al, 2012[197]	Substantial day-to-day pain variability was observed; higher levels of depression significantly predicted greater pain variability, but not anxiety.
Schneider et al, 2018[198]	Different pain components were identified - mean pain, mean amplitude of pain shifts, mean persistence of pain states (derived from the transition probabilities of pain states during the waking hours) and dominance of pain states.
Smith et al, 2008[199]	Anxiety and depression were significantly related to elevations in current and next week pain, although the effects were double for anxiety.
Van Dartel et al, 2013[200]	Pain and fatigue show synchronous monthly fluctuations and no temporal relationship with a time lag.
Ward 1994[201]	Depression, not mood, was found to have higher influence in variation in changes in pain and global arthritis status.

Ward 1993[202]	Pain and functional disability were significantly related to changes in global arthritis status over time, but did not vary with the duration of RA.
Wolfe et al, 2007[203]	The 0-10 VAS pain scale was better correlated with all RA clinical variables than the SF-36 scale; pain increased marginally with the duration of RA; pain decreased with age; higher pain was associated with ethnic minorities, female gender, and low education.

Anti-CCP: anti-citrullinated cyclic peptides; DAS28-P: 28-joint disease activity score (proportion of the DAS28 of patient global assessment); PGA: patient global assessment; RA: rheumatoid arthritis; SF-36: short form 36 items; SJCs: swollen joint counts; TJCs: tender joint counts; VAS: visual analogue score

Supplementary File

[Longitudinal Characteristics of Pain in Rheumatoid Arthritis: A Scoping Review]

Search database: MEDLINE, EMBASE, PSYCINFO

MEDLINE

- Database(s): Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions 1946 to May 27, 2022
- Search Strategy:

#	Searches	Results
1	exp Arthritis, Rheumatoid/	121505
2	((rheumatoid or rheumatic) adj2 (arthritis or disease* or disorder* or condition*)).mp.	186830
3	1 or 2	200807
4	exp Pain/	436086
5	pain.mp.	811803
6	4 or 5	894867
7	((trajector* or course or evolution or temporal or pattern* or longitudinal or variability or fluctuat* or progress* or chang*) adj4 pain).mp.	26471
8	6 and 7	26471
9	3 and 8	498
10	limit 9 to english language	443
11	limit 10 to (case reports or guideline or "review" or "systematic review")	114
12	10 not 11	329

EMBASE

- Database(s): Embase 1980 to 2022 Week 21
- Search Strategy:

#	Searches	Results
1	Arthritis, Rheumatoid.mp. or exp rheumatoid arthritis/	212870

2	((rheumatoid or rheumatic) adj2 (arthritis or disease* or disorder* or condition*)).mp.	278624
3	1 or 2	282206
4	exp pain/ or pain.mp. or exp chronic inflammatory pain/ or exp inflammatory pain/ or exp musculoskeletal pain/	1788872
5	((trajector* or course or evolution or temporal or pattern* or longitudinal or variability or fluctuat* or progress* or chang*) adj4 pain).mp.	40654
6	4 and 5	40654
7	3 and 6	1240
8	limit 7 to english language	1151
9	limit 8 to (meta analysis or "systematic review")	66
10	8 not 9	1085
11	limit 10 to "review"	94
12	10 not 11	991

PSYCINFO

- Database(s): APA PsycInfo 1806 to May Week 4 2022
- Search Strategy

#	Searches	Results
1	rheumatoid arthritis.mp. or exp Rheumatoid Arthritis/	3011
2	((rheumatoid or rheumatic) adj2 (arthritis or disease* or disorder* or condition*)).mp.	3620
3	1 or 2	3620
4	exp Pain/ or pain.mp.	123116
5	((trajector* or course or evolution or temporal or pattern* or longitudinal or variability or fluctuat* or progress* or chang*) adj4 pain).mp.	5897
6	4 and 5	5897
7	3 and 6	87
8	limit 7 to english language	83

9	limit 8 to ("0800 literature review" or "0830 systematic review" or 1200 meta analysis or 1600 qualitative study)	5
10	8 not 9	78

Table 1: Characteristics of included studies

Table 1	Country	Study design	Sample no.	Sample population	Gender	Mean age (years)	Mean disease/symptom duration (years)
Affleck et al, 1991	US	Prospective cohort study	47	Patients with RA recruited from the facility rheumatology practice of a university health center and community rheumatology practices in the same geographic area	76.6% female	52.77 +/- 12.79	9.06 +/- 7.66
Affleck et al, 1992	US	Prospective cohort study	75	Patients with RA recruited from the facility rheumatology practice of a university health center and community rheumatology practices in the same geographic area	71% female	52.7 (SD 12.5)	8.9 (SD 7.7)
Bailly et al, 2018 (abstract)	France	Prospective cohort study	165 (86 with RA)	Patients from ActConnect, a prospective cohort study	81% female (RA)	48.7 (SD 12.7 for RA)	8 (SD 8.8 for RA)
Cruise et al, 1996	US	Prospective cohort study	18/35	Voluntary patients with RA recruited through a local physician's office	25 F and 10 M	52.4 (SD 12.3)	Not specified
Drouet et al, 2022 (letter to the Editor)	UK	Prospective cohort study	165 (86 with RA)	Patients with RA recruited through a local physician's office (voluntary)	63% female	Median age 45	Median 8 years
Eberhard et al, 2021	Sweden	Prospective cohort study	232	Patients with RA recruited through a local physician's office (voluntary)	70.3% female	60.5 (SD 14.6)	Median 7 months (IQR 5-10)
Finan et al, 2010	US	Prospective cohort study	170	Recruitment was conducted for patients with RA in Phoenix, Arizona metropolitan area at local health fairs, physicians' offices, the Phoenix Veterans Affairs Hospital and through the Arthritis Foundation; participants received $90 for their participation	125 F 45 M	55.2 (SD 13.3)	Not specified
Gandrup et al, 2022	UK	Prospective cohort study	20	Patients were recruited from the rheumatology outpatient clinic at a single hospital site (Salford Royal NHS Foundation Trust, UK) in 2016	70% female	Median age 58.5 (IQR 48-64)	Not specified
Harries et al, 2014 (abstract)	UK	Prospective cohort study	1236	Patients with early RA from the inception cohort study, the Early Rheumatoid Arthritis Network (ERAN) from the UK and Eire	Unknown	Not specified	Not specified
Kumaradev et al, 2022	France	Prospective cohort study	810	Participants from two ongoing prospective French multicentric cohorts in a setting of universally accessible health-care: ESPOIR (Etude et Suivi des Polyarthrites Indifferenciees Recentes) and DESIR (DEvenir des Spondyloarthropathies Indifferenciees Recentes)	23.2% male	48.1 (SD 12.2) for RA	0.6 (SD 0.7)
Lotsch et al, 2020	Sweden	Retrospective cohort study	288	Participants based on the registry on the Swedish Epidemiological Investigation of RA with data collected between 2011 to 2016	209/288 women	52.2 (SD 12.3)	Not specified
McWilliams et al, 2019	UK	Prospective cohort study	683 (ERAN), 7090 (BSRBR - biologics), 1720 (BSRBR - non-biologics)	The ERAN inception cohort collected data from outpatient clinics in the UK and Eire with diagnosis of RA by a consultant rheumatologist. The BSRBR cohorts collected data from outpatient clinics in the UK with active RA. The BSRBR biologics cohort recruited people who were starting biologics as part of their routine care, having failed to respond adequately to other DMARDs. The BSRBR non-biologics cohort recruited people with RA who were using non-biologic DMARDs.	66% (ERAN), 77% (BSRBR - biologics), 75% (BSRBR - non-biologics)	57 (ERAN), 57 (BSRBR - biologics), 61 (BSRBR - non-biologics)	0.8 [SD 1.3] (ERAN), 13 [SD 10] (BSRBR - biologics), 10 [SD 11] (BSRBR - non-biologics)
Odegard et al, 2007	Norway	Prospective cohort study	149	Patients from the European Research on Incapacitating Diseases and Social Support (EURIDISS)	76% female	50.2 (SD 12.5)	2.2 (SD 1.2)
Roche et al, 2003	Australia	Prospective cohort study	120	Patients with RA attending an outpatient clinic	72% female	58.75 +/- 12.58	15.27 +/- 10.09
Sandberg et al, 2014 (abstract)	Sweden	Prospective cohort study	1802	Patients from a Swedish population-based case-control cohort of EIRA: Epidemiological Investigation of Rheumatoid Arthritis and follow-up in the Swedish Rheumatology Quality register	Not specified	Not specified	Not specified
Schneider et al, 2012	US	Prospective cohort study	106 (with rheumatic diseases) and 29% with RA	Patients recruited from two local community rheumatology practices	86% female	56 (SD 11)	Not specified
Schneider et al, 2018	US	Prospective cohort study		Patients from community rheumatology practices	84% female	57 (SD 13.12)	Not specified
Smith et al, 2008	US	Prospective cohort study	170 (82 with RA)	Voluntary patients recruited via methods such as newspaper ads, mailing to the Arthritis Foundation member and through rheumatology clinics	Not specified	63.8 (SD 7.3) and 62.1 (SD 7.3) for RA	Not specified
Van Dartel et al, 2013	The Netherlands	Prospective cohort study	228	Patients with RA attending the outpatient clinics of the Radboud University Nijmegen Medical Centre	64% female	56.7 (SD 10.6)	Median 10 (IQR 6-17)
Ward 1994	US	Prospective cohort study	24	Patients recruited from the Stanford University Medical Centre rheumatology clinics, from participants in local Arthritis Self-Management Courses, and from the community	22 F 2 M	Median age 46 (IQR 28-73)	Median 3 (IQR 0.2-20)
Ward 1993	US	Prospective cohort study	305	Participants from the Stanford Outcomes in Rheumatoid Arthritis study, a prospective, longitudinal study of health status and outcomes in a community based sample of patients with RA; live in Santa Clara County, California; recruited by advertisement between 1978 and 1981	83% female	53.9 +/- 13.8	14.1 +/- 9.0
Wolfe et al, 2007	US	Prospective cohort study	12090	Patients from the National Data Bank for Rheumatic Diseases longitudinal study of RA outcomes	23.17% male	60.09 (SD 13.61)	14.27 (SD 11.12)

Table 2: Description of methods and results for the included studies

Author, year	Pain scale	Frequency/timepoints	Study duration	Early/established RA	RA treatment	Methods for longitudinal pain analysis	Results - pain outcome	Analysis for covariates/predictors	Results for analysis of covariates/predictors
Affleck et al, 1991	VAS (no pain - very severe pain)	Daily	75 days	Established RA	Not specified	1. Measures of distribution of daily pain for each pain series: mean, standard deviation, and skew and outliers using box-and-whisker diagram 2. Measures of predictability: to identify linear trend, by regressing pain scores on recording day; first-order autocorrelation in each pain series (the correlation between pain on successive days); the maximum number of consecutive lagged days over which the autocorrelation function remained significantly different from zero (the duration of predictability)	1. Mean of the average daily pain score for each subject was 13.31 (SD = 9.77) with mean standard deviation of 4.05 (SD 2.18) and mean skew of 0.73 (SD 0.80) 2. Common distribution of positive skew (52.8%); only 2 pain series demonstrated statistically significant negative skew (higher frequency of painless days) 3. Trending series: 11/28 with downward progression (less pain) and 17/28 with upward progression (more pain) over the 75 days 4. Significant predictability of pain scores from one day to the next (significant first-order autocorrelations) with pain scores being highly correlated across successive days in most pain series (87.2%) and remains significant after recalculation of the first-order autocorrelations for detrended pain series (detrending was done by replacing each day's pain score with its residual from the regression of daily pain on recording day number, which then centers each series at zero) 5. Outlier analysis (painful and painless days): 51.1% with at least 1 painful day and no painless days, 2.2% with at least 1 painless day and no painful days and 10.6% had both painful and painless days; on 40 out of 75 occasions of reported isolated painful day, pain was rising during the previous 2 or more days (M = 2.8 preceding days)	Correlation with disease activity, physical disability and depression	Patients with more active disease, more disability and more depression had higher levels of mean daily joint pain, accounting for 36% of variance in mean daily pain [F(3,42)=5.92], with both depression (beta = 0.28) and disability (beta = 0.38) being unique predictors of average daily pain levels. The pain series in those with less active disease trended less sharply, contained more outlying painful days and longer episodes of comparatively severe pain, and had lower autocorrelations across successive days.
Affleck et al, 1992	VAS (no pain - very severe pain) from the Rapid Assessment of Disease Activity in Rheumatology	Daily	75 days	Established RA	Not specified	1. Descriptive analysis for the daily pain coping and correlation analysis between each coping components 2. Descriptive analyses of the mean and linear trends of the pain and mood time series and subsequent regression of the daily score on its position in the 75-day recording period; correlation analysis between average daily pain and average daily mood 3. Multivariate analyses using hierarchical regression modelling to examine the ability of daily coping scores to predict daily pain and mood (independent of pain) and trends in these variables while controlling for other significant correlates of pain and mood	Mean daily pain was 12.95 (SD 9.01), equivalent to 4 extremely painful joints or 6 moderately painful joints; mean beta weight for the daily pain trend was -0.03 (masking the individual differences in the direction of trend); 23 individuals had a significant downward (improving) trend over time and a nother 20 individuals with a significant upward (worsening) trend over time	1. Daily pain coping - 7 coping strategies (yes/no) and with each coping strategy, to identify specific behaviours used to take (direct action/relax/distraction); 2. Daily mood recording at night (from the Profile of Mood States-B - POMS-B)	Mean coping reports of 109.8 (SD 122.8); 40% of participants reported an average of at least 1 coping strategy/day; 10% used only 1 type of coping strategy while 24% used all 7 possible forms; a higher number of coping efforts were reported by women and individuals with greater disability and higher neuroticism, with gender (beta = 0.29) and disability (beta = 0.30) were independent predictors of total coping; women were more likely to used a greater number of different forms of coping and to seek emotional support; those with more disability were more likely to seek spiritual comfort; those with higher neuroticism were more likely to express emotions and were less likely to use relaxation as coping strategy; mean daily mood scores was 13.08 (SD 7.69), indicating more positive mood on most days; mean beta weight for the mood trend was 0.08 with 21 individuals trended significantly upward (improving mood) and 11 individuals trended significantly downwards (worsening mood); correlation between average daily pain and average daily mood was -0.37; daily pain was more intense in tohse with greater disability and who preceived less control over their daily pain; daily mood was more positive in those with lower neuroticism and with greater control over their daily pain; more coping overall was seen in those with more intense pain; specifically, those using relaxation as a coping strategy had less pain and those using emotional expression and seeking spiritual comfort had more intense pain; those with more coping had a more pronounced trend of improving pain and mood over time, also in those with greater control over their daily pain and those who used a greater number of different coping strategies; disability and pain control accounted for 30% of variance of daily pain intensity; at low levels of pain, greater use of coping strategies related to a more positive mood but, at high levels of pain, related to less positive mood
Bailly et al, 2018 (abstract)	NRS (0-10)	Weekly	3 months	Established RA	44% with RA on a biologic DMARD	Mean of the inter-assessment differences of NRS (average of absolute change - AAC) of pain and disease activity for both RA and axSpA and were compared by t-test; with high variability defined in the upper tercile of AAC	Mean pain at baseline of 2.90 (SD 2.36) with low AAC of pain in RA of 1.02 [SD 0.74] - low fluctuations of pain, which was around 1 point on a scale of 0-10	1. Pearson's correlation to evaluate the correlation between variability of pain and of self-reported disease activity; 2. univariate and multivariate logistic regression comparing patients with high vs low variability of pain (without missing data imputation), including self-reported flare over follow-up	Correlation between AAC of pain and activity was 72% in RA; in multivariate analyses, self-reported flares were the only determinants of pain AAC with OR of 2.25 [1.27-4.38] for RA; fluctuations in pain were highly correlated to fluctuations in PGA of disease activity, indicating significant overall for these 2 patient-reported outcomes, with self-reported flares contributing to the fluctuations in pain (confirming the validity of self-reported flares)
Cruise et al, 1996	7-point scale (0-6 with 0 = not at all, 3 = moderate, and 6 = extremely)	7 times daily (0800 - 2100)	7 days	Not specified	Not specified	Repeated measures ANOVAs	Mean pain rating remains stable over the 7-day period (between scale of 1.5-2.0)	Correlation with mood states (happy, depressed/blue, joyful, unhappy, angry/hostile, enjoyment/fun, frustrated, worried/anxious, pleased, energetic, relaxed, fatigued, alert) and frequency of significant events; repeated measures of ANOVAs for effect of study day on mood and significant events	Fatigue was positively correlated with pain; positive mood items all correlated negatively with pain intensity, but only alertness and energy were significantly negatively related to pain' no reactive effects were observed in ratings of pain, positive mood, negative mood or frequency of significant events, also in the analysis between high- and low-pain groups
Drouet et al, 2022 (letter to the Editor)	NRS (0-10)	Weekly	12 weeks	Established RA	53% on a biologic DMARD (pooled result)	Description of flucutations of pain over time with variability defined as an absolute variation of more than 2 points between 2 consecutive assessments	Aggregated results 1. Pain remained stable over time at group level (P = 0.54); 2. At within-individual level, there was no pain fluctuation in 31 (19% patients; 22% in RA); 3. 60 (36%) patients with high pain fluctuations had fluctuations of >2 points in at least 4 of 11 intervals (34% in RA)	Univariable and multivariable logistic regression to compare the tertile of patients with most frequent fluctuations with the less fluctuating ones - baseline characteristics	Aggregated results - Lower age, lower disease duration, disease active at baseline, higher BASDAI, higher pain and PGA at baseline, symptomatic treatment intake and the absence of bDMARDs intake were significantly associated with frequent fluctuation; in multivariable analysis, female gender (OR 2.94 [95% CI 1.25-7.35]) and higher PGA at baseline (OR 2.00 [95% CI 1.36-3.07] were significantly associated with frequent pain fluctuations; also a non-significant trend of more pain fluctuations in those with active disease at baseline (OR 2.10 [95% CI 0.98-4.57]
Eberhard et al, 2021	VAS (0-100mm)	5 irregular timepoints (at inclusion, 6 months, and 1, 2 and 5 years)	5 years	Early RA	61.5% were on methotrexate at 5 years; during the 5-year period, 17% were at some point treated with a biologic DMARD.	1. Descriptive statistics for development of pain over time; 2. Change in pain between every visit using paired t-test; 3. Estimation of mean differences in pain over time and differences in change of pain per month	Mean VAS pain was 41.2 at baseline and decreased significantly to 32.3 at the 6-month visit, but remained unchanged for the rest of the follow-up period; mean change in VAS pain at inclusion to 6 months was -9.2, and no significant change in pain was observed between the follow-up visits; 49.1% had unacceptable pain at inclusion and lower proportion of 30.1% during hte first year; 20.2% at inclusion had unacceptable pain with low inflammation and this proportion did not change significantly during the 5-year follow-up; 35% had unacceptable pain (= VAS >40mm) 5 years after study inclusion	Univariate and multivariate logistic regression analysis used to assess potential baseline predictors of unacceptable pain, and of unacceptable pain with low inflammation; baseline predictors of pain over time were assessed using mixed model analysis, using all VAS pain values at inclusion and follow-ups at 6 months, 1, 2 and 5 years	Baseline predictors of unaccpetable pain with low inflammation at 1 year were higher VAS pain, lower age and lower ESR; negative associations between CRP in the highest quartile at baseline and unacceptable pain with low inflammation was seen; female gender at baseline was associated for those with unacceptable and low inflammation at 2 years; at 6-month visit, there was a negative association with baseline erosion; in multivariate analysis, lower age was associated with unacceptable pain with low inflammation at 2 years (also similar trend at 1 year); at 5-year follow-up, 40 (22.5%) had unacceptable pain with low inflammation, with negative anti-CCP being the only significant baseline predictor of this state in both univariate and multivariate analysis; in multivariate analysis, higher baseline VAS PGA was a significant predictor, with a similar trend for lower SJC at baseline; between 6 months to 5 years of follow-up, 10-15% had unacceptable pain with high inflammation (significant baseline predictors at 1- and 2-year follow-up include seropositivity, high inflammatory markers, high DAS28 and severe PRO, but no significant predictors identified at 5 years); in mixed model analysis, baseline predictors of increased pain over time include higher PGA (strongest), HAQ, DAS28, TJC28, ESR and CRP; higher SJC at baseline had higher pain level at baseline, but no significant difference in pain over time; no association with seropositivity; those with worse PROs and disease activity measures at baseline had greater reductions in pain over time and those with higher grip force at baseline has less reductions in pain over time; older patients had less pain at baseline and over time, but also less likely to experience reduced pain during follow-up
Finan et al, 2010	NRS (101-point)	Daily	29 days	Not specified	Not specified	Non-linear damped oscillator model used to examine the nature of fluctuations across time in pain prediction accuracy	Pain prediction accuracy oscillated over time; the oscillation amplitude was larger at the start of diary and decreased over the time series, indicating damping toward more accurate predictions (people become more accurate over time in predicting their pain as they gain experience to do so)	Associations of affect (positive/negative) and degee of pain control with the oscillation of pain over time	The oscillations for individuals with lower negative affect and higher pain control damped more quickly than the oscillations for their counterparts, indicating more accurate pain prediction process in the chronically ill; positive affect did not differentiate the damping pattern, but within each oscillation cycle, individuals with higher positive affect spent more time making inaccurate predictions than their counterparts

Table 2: Description of methods and results for the included studies

Gandrup et al, 2022	NRS (0-10)	Daily	85 days	Not specified	Not specified	Descriptive summary of mean pain	1. Mean pain during flare weeks was 4.4 (SD 1.8) and during non-flare weeks was 3.6 (SD 1.9) 2. Standard deviation of pain symptom in the week before flare reporting was 1 (SD 0.4) during flare weeks and 0.7 (SD 0.4) during non-flare weeks 3. The slope of pain symptom score in the week before flare reporting was 0.12 (SD 0.19) during flare weeks and -0.01 (SD 0.14) during non-flare weeks.	Mixed effects logistic regression to quantify associations between flare weeks and symptom scores; multivariate modelling for pain symptoms and flare	Univariate mixed effects models showed higher mean and steeper upward slopes in symptom scores in the week preceding the flare increased the likelihood of flare occurrence, but less prominent association with variability; in the multivariate modelling using pain symptom, mean scores (OR 1.83 [95%CI 1.15-2.97]) and variability of pain (OR 3.12 [95%CI 1.07, 9.13]) were associated with higher odds of flare
Harries et al, 2014 (abstract)	SF-36 bodily pain	Unknown (completed at each follow-up visit	3 years	Early RA	Not specified	Descriptive summary of mean pain	Pain improved from baseline to 1 year (median 41 [IQR 22-62] and median 51 [IQR 31-72] respectively) and then remained constant afterwards	Generalised estimating equation (GEE) analyses, adjusting for confounders, to examine the associations of pain with demographic and clinical characteristics at baseline and each follow-up	DAS28-P at baseline had median of 0.42 (IQR 0.35-0.51) and did not change substantially during the 3-year follow-up; initial GEE analysis showed high DAS28-P was consistently associated with worse pain throughout the follow-up and that high DAS28-P significantly predicted the next year's pain; disability (HAQ), fatigue (SF-36 - vitality) and current pain may predict future pain better than other well-established RA disease markers
Kumaradev et al, 2022	SF-36 bodily pain and VAS (0-100) for joint pain when mobilising and when at rest	Biannually in the first 2 years, then annually	Up to 10 years	Early RA	Biologic DMARD naïve at inclusion; at baseline, 90.9% on NSAIDs, 19.7% on cDMARDs, 6.9% on bDMARDs and 67.8% on analgesics	Descriptive summary of mean pain	For ESPOIR - RA cohort, the mean SF-36 bodily pain at baseline was 62.2 (SD 20.4) with mean VAS of 54.9 (SD 25.8) when mobilising the joints and mean VAS of 37.0 (SD 27.5) when at rest.	Linear mixed models were used to characterise differences in pain evolution as a function of age (tertiles), sex, etnicity, education, marital and professional status, after accounting for disease-related, treatment, lifestyle and health factors	With the transition from early disease (=<6 months for RA and =<3 years for SpA) to long s-tanding disease, differences in pain evolution emerged as a function of age, sex and ethnicity in RA; being younger age, males, and Caucasians exhibited lower pain in the latter phases of both diseases; highly educated participants (beta = -3.8 in RA) for both diseases presented with low pain early in the disease, with no changes throughout disease course
Lotsch et al, 2020	VAS (0-100mm)	At baseline, 3, 6, 12, 24, 36, 48 and 60 months	5 years	Early RA	Not specified	Unsupervised machine learning clustering to identify homogenous subgroups of pain trajectories	3 distinct subgroups identified: low-, medium- and high-persistent pain intensity	Supervised machine learning implemented as random forests followed by computed ABC analysis-based item categorisation was used to identify predictive parameters among the demographics, patient-rated, and objective clinical factors	With 3-month data (alogrithms-trained), patient global assessment and health assessment questionnaire provided pain group assignment at a balanced accuracy of 70%; when examining the objective clinical parameters alone on an alogrithm-trained 3-month data, this provide a balanced accuracy of 59% (using disease severity, SJC and TJC)
McWilliams et al, 2019	SF-36 bodily pain	Unknown (completed at each follow-up visit	3 years	Both early RA (ERAN) and established RA (BSRBR-biologics and BSRBR-non-biologics)	The BSRBR biologics cohort recruited people on biologics, having failed to respond adequately to other DMARDs. The BSRBR non-biologics cohort recruited people who were using non-biologic DMARDs.	Descriptive summary of mean pain and clustering analysis of the pain trajectories	Mean SF-36 bodily pain scores in each cohort improved but remained throughout 3 years of follow-up of >1 standard deviation worse than the UK general population; discrete persistent pain (59-79%) and resolving pain (19-27%) were identified in each cohort, but in ERAN, a third trajectory of persistently low pain (23%) was identified.	Logistic regression to compare baseline predictor variables between trajectories; the role of inflammation was examined in a subgroup analysis of people with normal levels of inflammatory markers after 3 years	In people with normal inflammatory markers after 3 years, 65% were found to follow a persistent pain trajectory; in the logistic regression analysis, greater disability (adjusted OR = 2.3-2.5 per unit baseline HAQ score) and smoking history (adjusted OR = 1.6-1.8) were risk factors for persistent pain trajectories in each cohort
Odegard et al, 2007	VAS (0-100mm)	5 irregular timepoints (at baseline, and after 1, 2, 5 and 10 years)	10 years	Established RA	At baseline, 56% had DMARDs, 25% had prednisolone and 52% used NSAIDs. 12% were on TNF inhibitors (infliximab or etanercept) either as monotherapy (5%) or in combination with methotrexate (7%) . 19% never used DMARDs.	Descriptive summary of mean pain and also repeated measures analyses of variance to explore the effect of time on pain	At various timepoints of pain measurement, 30% had VAS >=40mm. Mean VAS pain score was 31.4 (SD 24.5) for complete follow-up participants.	Repeated measures analyses of variance to explore the effect of time on measures of outcome including anxiety, disease activity, physical function and depression; mixed model analyses to identify factors that were longitudinally associated with pain, depression and anxiety	5-13% had an AIMS depression score of >=4.0 and 20-30% had an AIMS anxiety of >=4.0; the pain level was explained by anxiety, disease activity, physical function and female gender; the depression was explained by high disease activity; the anxiety was explained by low disease activity and depression
Roche et al, 2003	VAS (0-10)	3 timepoints (at baseline, at 63 months, and at 77 months)	77 months	Established RA	Not specified	Binomial categories for the Pain Rating Index (PRI) scores into mild-moderate pain or severe pain; ANOVA for independent samples or chi-square (categorical data) for differences between reassessed subjects and those lost to follow-up; One-way repeated measures ANOVA for differences across the 3 timepoints (continuous data)	The average pain VAS and PRI scores at baseline were moderate and remained unchanged over time	Friedman's test to examine the differences in the categorical measures of medication, disease activity and functional capacity; Chi-square and ANOVA were used to test for differences in the average SU and RW	Significant pain sensory components found include pressure and constriction; pain-related affect was described with adjectives suggesting positive psychological adaptation to pain
Sandberg et al, 2014 (abstract)	VAS (0-100mm)	6 timepoints (at baseline, 3 months, 6 months, 12 months, 2 years, and 5 years)	5 years	Early RA	Not specified	Limited in an abstract format	Not specified	Correlation analyses between pain and DAS28, ESR, CRP, TJC, SJC and PGA	At baseline, pain was correlated to both inflammation (ESR) and clinical arthritis, but after 3 months, there was loss of correlation between pain and objective inflammation (both ESR and CRP); sustained and stable correlation between pain and SJC/TJC for the whole study period
Schneider et al, 2012	VAS (101-point)	Daily	28 days	Not specified	Not specified	Multilevel modelling analyses	Substantial day-to-day pain variability with significant inter-individual differences. The total variance was 34% for within individual day-to-day pain fluctuations.	Multilevel modelling analyses	Higher levels of depression, happiness and frustration significantly predicted greater variability in pain; lower self-efficacy was associated with more variability in patients' daily satisfaction with accomplishments and in the quality of their day; greater pain catastrophising and higher depression predicted more variability in interference with social relationships; anxiety was not significantly associated with day-to-day variability
Schneider et al, 2018	VAS (0-10)	Daily	3 months	Not specified	Some patients enrolled in the study were about to start a change in their pain treatment (i.e., starting a new treatment, adding a new treatment to the current regimen, switching to a different treatment, or increasing the current treatment dose) after the baseline assessment.	Markov regime switching dynamic regression model	Mean VAS (0-10) pain was 3.84 at Month 1 and 3.39 at Month 3; mean amplitude of pain shifts was 3.47 at Month 1 and 3.32 at Month 3 (approximately 1/3 of the full scale range); mean persistence of pain states (derived from the transition probabilities of pain states during the waking hours) was 10.3 hours at Month 1 and 9.9 hours at Month 3; for the dominance of pain states, lower pain states tended to be more enduring than higher pain states; in Month 1, higher states were on average 1.44 hours shorter than lower pain states and in Month 3, higher pain states were on average 4.85 hours shorter than lower pain states; moderate positive associations were seen between patients' average pain level and the persistence of states at each assessment period (Month 1: r = 0.38, Month 3: r = 0.26), suggesting patients with overall higher pain levels experienced more enduring pain states; the persistence and amplitude measures were negatively correlated at each assessment period (Month 1: r = -0.36, Month 3: r = -0.37), suggesting patients with more enduring pain states tended to experience less pronounced pain shifts; test-retest correlations showed moderate-high stability in average pain (r = 0.72) and amplitude (r = 0.75) measures, and low-moderate stability in persistence (r = 0.23) and dominance (r = 0.47) measures over the 3-month period	To examine cross-sectional and longitudinal associations of the regime-switching pain measures with patients' physical functioning and emotional health outcomes; to assess if the new measure contribute to understanding patients' retrospective judgements of their pain and to their global impressions of change over time	The mean BDI depression score was 10.67 (SD 7.09); the mean STAI anxiety score was 37.33 (SD 9.64); the mean SF-36v2 physical component score was 34.81 (SD 8.93) and the mean mental component score was 47.98 (SD 9.82); at the beginning, higher Average pain levels were signficantly associated with higher depression and anxiety, and with lower mental and physical health scores (9.3%-13.9% of the variance) and in month 3, the Persistence of pain states was associated with depression and general mental health (3.6%-5.2% of the variance); for patients with depression scores in the minimal range (BDI =<13), a given pain state lasted between 7.7-10.9 hours before transitioning to other pain state, and those with moderate depression (BDI scores of 20-28), the pain states lasted longer between 13.9-18.5 hours; increases in Average pain were significantly associated with worsening in each of the health outcomes (16-32% of the variance) in Step 1 of the regression analyses; for negative affect, increases in Persistence of pain states were associated with increases in negative emotional states; increases in the Dominance of higher pain states were associated with increases in pain interference, physical functioning limitations and ADL difficulties; with regards to global overall impressions of change, about 50% reported that they had pain improvement over the 3-month period, 24! reported pain as unchanged and about 25% reported worsening pain over time
Smith et al, 2008	Questionnaire 0-100 of average pain	Weekly	11 weeks	Not specified	Not specified	Multilevel modelling analyses	Mean pain of 45.32 (SD 22.10) for RA	To examine the effects of anxiety and depression (measured at baseline) on weekly pain reports; to examine the associations with interpersonal stress and positive and negative affect	When anxiety and depression were examined separately in the multilevel analyses, anxiety and depression were both related to elevations in current and next week pain, with anxiety having the largest effect size; both anxiety and depression were indirectly related to current pain through negative and positive affect and depression interacted with stress to predict current pain in RA; when both anxiety and depression were examined simultaneously, anxiety alone was still related to elevations in current and next week pain and was indirectly related to current pain through negative affect; depression alone was indirectly related to current pain through positive affect
Van Dartel et al, 2013	NRS (0-10)	Monthly	12 months	Established RA	At baseline, 41.4% received DMARD monotherapy, most often with methotrexate, sulfasalazine and azathioprine. 12% received DMARD combination therapy, usually with methotrexate. 35.4% received bDMARD and all of them received TNF-inhibitors. 4% stopped a biologic agent during the study and 6% started a bDMARD during the study. 13% received prednisolone.	Longitudinal mixed regression analyses (time-lag model)	Mean pain score at baseline was 35.23 +/- 19.82 and the mean patient-averaged pain score over 1 year was 36.4 +/- 18.3.	Longitudinal regression analyses and time-lag models to examine the relationship between fatigue and pain and change over time; to examine the associations differed by sex and age (effect modification)	A significant positive relationship between fatigue and pain during the same month (beta = 2.04; 95%CI 1.82, 2.27) in the longitudinal regression analysis; in the time-lag models, there was no significant association between changes in pain and changes in fatigue; the associations between change in pain score and change in faitgue score were not significantly different for men and women, nor did they vary with age

Table 2: Description of methods and results for the included studies

Ward 1994	VAS (15cm) with 0 = no pain and 3 = very severe pain	Every 2 weeks	60 weeks	Established RA	Not specified	Pooled time series regression analysis	Initial median pain was 1.2 with maximal percent change from the initial value of 92%.	To determine the extent to which mood and depression may confound self-reported measures of functional disability, pain and global arthritis (with the effects of these measures were being controlled in the equation)	Mood (POMS-B) scale explained 2.0% or less of the variation in all longitudinal measures , after controlling for the effects of the clinical measures of arthritis activity; depression (CES-D) explained <2.0% of the variation in changes in functional disability, but explained 6.0% and 8.0% of the variation in changes in pain and global arthritis status, respectively
Ward 1993	HAQ VAS (15cm) with 0 = no pain and 100 = very severe pain	Every 6 months	Up to 9.5 years	Established RA	Not specified	Pooled time series regression analysis	Median pain score at baseline was 1.1 (with mean of 1.2 +/- 0.8).	To determine the relative importance ascribed to factors of pain and functional disability by examining the relationship of levels of both pain and functional disability to self-reported global arthritis status by using pooled time series analyses	Median level of functional disability was 1.25 (IQR 0-3) and median level of global arthritis status was 35 (IQR 0-100) and together with the median level of pain, these results indicate moderate arthritis severity; both pain and functionl disability were significantly related to changes in global arthritis statu over time; pain seems to be an important feature in non-white cohort and disability was more important for men; duration of RA did not influence the relative importance of pain and disability
Wolfe et al, 2007	VAS (0-10)	Every 6 months	Not specified	Established RA	Not specified	Descriptive summary of mean pain	Mean VAS pain was 3.4 (SD 2.8) with the best cutpoint for an 'acceptable' level of pain of =<2.0; pain levels were almost constant over RA duration	Regression analyses (GEE) - univariable and multivariable	Pain increased slightly with the duration of RA, 0.03 (95% CI 0.02-0.03) and decreased with age, 0.01 (95% CI 0.01-1.02) units per year; pain levels increased in women, ethnic minorities, smokers and those with less education

List of Abbreviations

AAC – average of absolute change
AIMS – Arthritis Impact Measurement Scales
ANOVA – analysis of variance
axSpA – axial spondyloarthritis
BASDAI – Bath Ankylosing Spondylitis Disease Activity Index
BDI – Beck Depression Inventory-II
bDMARDs – biologic disease modifying anti-rheumatic drugs
BSRBR – British Society for Rheumatology Biologic Registry
cDMARDs – conventional disease modifying anti-rheumatic drugs
CI – confidence interval
CRP – C-reactive protein
DAS28 – 28-joint disease activity score
DESIR – DEvenir des Spondyloarthropathies Indifferenciees Recentes
DMARDs – disease modifying anti-rheumatic drugs
EIRA – Epidemiological Investigation of Rheumatoid Arthritis
ERAN – Early Rheumatoid Arthritis Network
ESPOIR – Etude et Suivi des Polyarthrites Indifferenciees Recentes
ESR – erythrocyte sedimentation rate
EURIDISS – European Research on Incapacitating Diseases and Social Support
GEE – generalised estimating equation
HAQ – Health Assessment Questionnaire
IQR – interquartile range
mm – milimetre(s)
NHS – National Health Service
NRS – numerical rating scale
NSAIDs – non-steroidal anti-inflammatory drugs
OR – odds ratio
PGA – patient global assessment
POMS-B – Profile of Mood States-B
PRI – Pain Rating index
PRO – patient-reported outcome
RA – rheumatoid arthritis
RW – rank of word
SD – standard deviation
SF-36 – Short Form 36 items
SJC – swollen joint count
SpA – spondyloarthritis
STAI – Spielberger State-Trait Anxiety Scale (STAI)
SU – subclass use
TJC – tender joint count
TJC28 – 28-jpint tender joint count
TNF – tumour necrosis factor
US – United States
UK – United Kingdom
VAS – visual analogue scale

References [1-22]

1. Affleck G, Tennen H, Urrows S, Higgins P. Individual differences in the day-to-day experience of chronic pain: a prospective daily study of rheumatoid arthritis patients. Health Psychol. 1991;10(6):419-26. DOI: 10.1037//0278-6133.10.6.419.
2. Affleck G, Urrows S, Tennen H, Higgins P. Daily coping with pain from rheumatoid arthritis: patterns and correlates. Pain. 1992 Nov;51(2):221-9. DOI: 10.1016/0304-3959(92)90263-B.
3. Bailly F JC, Servy H, Fautrel B, Gossec L. Variability of pain levels is explained by self-reported disease acitvity in rheumatoid arthritis and axial spondyloarthritis: A 3-months prospective study of 165 patients. Annals of Rheumatic Diseases 2018;77:538.
4. Cruise CE, Broderick J, Porter L, Kaell A, Stone AA. Reactive effects of diary self-assessment in chronic pain patients. Pain. 1996 Oct;67(2-3):253-8. DOI: 10.1016/0304-3959(96)03125-9.
5. Drouet J, Gossec L, Jacquemin C, Fautrel B, Foltz V, Gandjbakhch F, et al. Fluctuation of pain is frequent in rheumatoid arthritis and axial spondyloarthritis: A 12 weeks prospective study of 165 patients. Joint Bone Spine. 2022 May;89(3):105306. DOI: 10.1016/j.jbspin.2021.105306.
6. Eberhard A, Bergman S, Mandl T, Olofsson T, Rydholm M, Jacobsson L, et al. Predictors of unacceptable pain with and without low inflammation over 5 years in early rheumatoid arthritis-an inception cohort study. Arthritis Res Ther. 2021 Jun 14;23(1):169. DOI: 10.1186/s13075-021-02550-7.
7. Finan PH HE, Amazeen PG, Butner J, Zautra AJ, Tennen H. Oscillations in daily pain prediction accuracy. Nonlinear dynamics, psychology and life sicences. 2010;14(1):27-46.
8. 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. DOI: 10.1093/rap/rkac021.
9. Harries L MD, Young A, Kiely PDW, Walsh DA. Prediction of future pain by DAS28-P in patients with early rheumatoid arthritis: The ERAN cohort. Rheumatology (UK). 2014;1:i71-i2.
10. Kumaradev S, Roux C, Sellam J, Perrot S, Pham T, Dugravot A, et al. Socio-demographic determinants in the evolution of pain in inflammatory rheumatic diseases: results from ESPOIR and DESIR cohorts. Rheumatology (Oxford). 2022 Apr 11;61(4):1496-509. DOI: 10.1093/rheumatology/keab562.
11. Lotsch J, Alfredsson L, Lampa J. Machine-learning-based knowledge discovery in rheumatoid arthritis-related registry data to identify predictors of persistent pain. Pain. 2020 Jan;161(1):114-26. DOI: 10.1097/j.pain.0000000000001693.
12. McWilliams DF, Dawson O, Young A, Kiely PDW, Ferguson E, Walsh DA. Discrete Trajectories of Resolving and Persistent Pain in People With Rheumatoid Arthritis Despite Undergoing Treatment for Inflammation: Results From Three UK Cohorts. J Pain. 2019 Jun;20(6):716-27. DOI: 10.1016/j.jpain.2019.01.001.
13. Odegard S, Finset A, Mowinckel P, Kvien TK, Uhlig T. Pain and psychological health status over a 10-year period in patients with recent onset rheumatoid arthritis. Ann Rheum Dis. 2007 Sep;66(9):1195-201. DOI: 10.1136/ard.2006.064287.

14. Roche PA, Klestov AC, Heim HM. Description of stable pain in rheumatoid arthritis: a 6 year study. J Rheumatol. 2003 Aug;30(8):1733-8.
15. Sandberg M SS, Altawil R, Alfredsson L, Klareskog L, Lampa J. The relationship between pain and inflammation in early rheumatoid arthritis: Long-term follow-up in the Swedish population-based EIRA and SRQ register. Scandinavian Journal of Rheumatology. 2014;127:4-5.
16. 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):813-22. DOI: 10.1016/j.pain.2012.01.001.
17. Schneider S, Junghaenel DU, Ono M, Stone AA. Temporal dynamics of pain: an application of regime-switching models to ecological momentary assessments in patients with rheumatic diseases. Pain. 2018 Jul;159(7):1346-58. DOI: 10.1097/j.pain.0000000000001215.
18. Smith BW, Zautra AJ. The effects of anxiety and depression on weekly pain in women with arthritis. Pain. 2008 Aug 31;138(2):354-61. DOI: 10.1016/j.pain.2008.01.008.
19. van Dartel SA, Repping-Wuts JW, van Hoogmoed D, Bleijenberg G, van Riel PL, Fransen J. Association between fatigue and pain in rheumatoid arthritis: does pain precede fatigue or does fatigue precede pain? Arthritis Care Res (Hoboken). 2013 Jun;65(6):862-9. DOI: 10.1002/acr.21932.
20. Ward MM. Are patient self-report measures of arthritis activity confounded by mood? A longitudinal study of patients with rheumatoid arthritis. J Rheumatol. 1994 Jun;21(6):1046-50.
21. Ward MM, Leigh JP. The relative importance of pain and functional disability to patients with rheumatoid arthritis. J Rheumatol. 1993 Sep;20(9):1494-9.
22. Wolfe F, Michaud K. Assessment of pain in rheumatoid arthritis: minimal clinically significant difference, predictors, and the effect of anti-tumor necrosis factor therapy. J Rheumatol. 2007 Aug;34(8):1674-83.

2.4 Summary

This chapter summarises the current literatures on the methods used in longitudinally characterising pain of adult populations in RA based on a literature search from inception to May 2022. Main findings from this scoping review, as outlined below, form the basis of conducting the research work in RA that I conducted and presented in the remainder of this thesis.

Main findings include:

- Heterogeneity in pain scales used and the frequency of pain measurements in RA exists, which is consistent with the paradigm shift in study designs and longitudinal data collections
- Despite the initial variability in pain level, stable pain trajectories in RA with continuing minimal pain fluctuation over time were seen, akin to the stable pain trajectories identified in OA studies
- The use of VAS and NRS pain scales used in the identified stable pain trajectories in RA necessitates the consideration of the ordinal nature of the scale used in the analysis, of which the interval between each pain level is assumed equal in order, although the magnitude of this interval is subjective to individuals experiencing the same pain level
- Distinct longitudinal analyses of pain in RA using clustering methods and novel dynamic time series methods were seen, which correlated with the research questions of interest in terms of evaluating either the extent or the magnitude of pain variability and the effect of time applied in the analyses
- Important predictors such as female gender, ethnicity, smoking history and education level were identified, suggesting the importance of considering these predictors in future pain trajectory research

Summary of Section 1

Section 1 presents the results of the scoping review in Chapter 2, which addresses the first research question in this thesis in terms of exploring the methods used and results generated from longitudinal characterisations of pain in RA. The evidence identified through this scoping review suggests that there is heterogeneity across studies in describing pain experience in RA. As pain in RA is highly variable and multidimensional, future longitudinal research in examining pain trajectory and its variability should incorporate important predictors of pain variability, standardised measures of pain, and appropriate frequency and intervals of pain measurements.

Specifically, this scoping review has highlighted the rationale behind pain trajectory analysis and the opportunities and challenges when interpreting the analysis findings. To address my thesis aims, this review has provided me the impetus to start my journey of learning the fundamentals of clinical epidemiology and intensive longitudinal analysis (Section 2 – Chapters 3 and 4), followed by the application of these research skills and knowledge when assessing pain trajectory in RA (Section 3 – Chapters 5, 6 and 7).

Section 2: Digitalised Health Data and The Application of Intensive Longitudinal Methods in Analysing Pain Using Mobile Health Data

Overview of Section 2

Section 2 of this thesis consists of two chapter, Chapters 3 and 4, which address the next research questions on digitalised health-related data and the application of intensive longitudinal methods in examining pain in RMDs.

Chapter 3 presents the narrative review of harnessing digitalised health-related information in creating real-world evidence in clinical care and research. Benefits and challenges of using such digitalised health-related data in modern medicine are discussed in this chapter.

Chapter 4 presents the research work done and research skills gained during my remote candidature period in the UK. This chapter presents the analyses of pain trajectory and day-to-day pain variability in inflammatory and non-inflammatory RMDs using a large smartphone study, *Cloudy with a Chance of Pain*, dataset. Intensive longitudinal methods were applied in the analyses. This chapter hones my research skills and improves my knowledge in examining pain trajectory in RMD, which is an instrumental step in conducting research work presented in Section 3 of this thesis.

Chapter 3: Digitalised Health Data in Providing Real-World Evidence: A Review

3.1 Preface

The narrative review presented in this chapter addresses the research question of harnessing digitalised health-related data, in terms of appraising the nature of these data collected and used, as well as the benefits and challenges around using such data in modern medicine and research.

This narrative review is presented in a manuscript format in this chapter, as it has been published in a peer-reviewed journal, Rheumatology, as part of the Real World Data: Special Section as:

Pisaniello, H.L., Dixon, W.G. What does digitalization hold for the creation of real-world evidence? Rheumatology (Oxford, England), 2020; 59(1):39-45.
doi:10.1093/rheumatology/kez068

The Statement of Authorship is included. To end, I present the chapter summary from this narrative review.

3.2 Statement of Authorship

Statement of Authorship

Title of paper	What does digitalization hold for the creation of real-world evidence?
Publication status	Published
Publication Details	Pisaniello, H.L., Dixon, W.G. What does digitalization hold for the creation of real-world evidence? Real World Data: special section. Rheumatology, 2020; 59:39-45 doi:10.1093/rheumatology/kez068 Impact Factor: 7.046

Principal Author

Name of principal author	Huai Leng Pisaniello
Contribution to the Paper	Contributed to the study design and conception, the development of literature search and literature review with critical appraisals conducted for articles cited in the manuscript. Contributed to the drafting and revision of the manuscript for publication.
Overall percentage (%)	75%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 8 April 2023

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. The candidate's stated contribution to the publication is accurate (as detailed above);
ii. Permission is granted for the candidate in include the publication in the thesis; and
iii. The sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	William G. Dixon
Contribution to the Paper	Contributed to the study design and conception. Supervised the development of literature search and literature review. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 6, 2023

3.3 Manuscript: What Does Digitalisation Hold for the Creation of Real-World Evidence?

RHEUMATOLOGY

Rheumatology 2020;59:39–45
doi:10.1093/rheumatology/kez068

Real World Data: special section

What does digitalization hold for the creation of real-world evidence?

Huai Leng Pisaniello[1,2] and William Gregory Dixon[1]

Abstract

Health-related information is increasingly being collected and stored digitally. These data, either structured or unstructured, are becoming the ubiquitous assets that might enable us to comprehensively map out a patient's health journey from an asymptomatic state of wellness to disease onset and its trajectory. These new data could provide rich real-world evidence for better clinical care and research, if they can be accessed, linked and analyzed—all of which are possible. In this review, these opportunities will be explored through a case vignette of a patient with OA, followed by discussion on how this digitalized real-world evidence could best be utilized, as well as the challenges of data access, quality and maintaining public trust.

Key words: real-world evidence, electronic health record, mobile app, accelerometers, digital data, unstructured data, data protection, osteoarthritis

Rheumatology key messages

- The volume and breadth of digital data contributing to real-world evidence is expanding.
- Digital data will allow researchers to answer questions that cannot currently be addressed.
- Real-world digital health data require robust data governance, sustainable public trust, data standardization and interoperability.

Introduction

The increased uptake of technology is changing our ability to observe and understand the onset, progression and outcome of disease in society. Information and communication is increasingly stored digitally. There has been an exponential expansion in stored data, from digital versions of traditional media like text, to images and videos, sensors, digital transactions and even digital traces of our interactions with technology [1, 2]. As we live our daily lives, vast amounts of information pertaining to our health and well-being are being recorded, including contact with health care systems. This includes our exposure to environmental and behavioral risk factors while living free from disease, the onset of symptoms and progress towards a clinical diagnosis, as well as the consequences and impact of living with a disease and its treatment. Digital data relevant to health are expanding from the more obvious and traditional, for example, notes held within a medical record, to the sometimes less apparent information captured about our everyday lives, as well as a wide array of data describing the environment in which we live. This digital archive of information is fragmented and scattered, sometimes unstructured, yet it has the potential to help us better understand diseases and their treatment and ultimately to improve the lives of future populations. This article will consider the wide range of data that exists, both traditional and novel, that might contribute to real-world evidence (RWE) about health and disease.

Spot the digital data sources

Let us consider a patient's health journey, starting from a pre-morbid state of wellness to the onset of disease symptoms, self-management and interactions with health care professionals, to treatment response and disease outcome. Through a working example, we will examine how a patient may seek help, information, support, guidance and treatment through this journey, with an eye on what digital data are captured.

[1]Arthritis Research UK Centre for Epidemiology, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK and [2]Department of Medicine, University of Adelaide, Adelaide, South Australia, Australia

Submitted 30 September 2018; accepted 7 February 2019

Correspondence to: William Gregory Dixon, Arthritis Research UK Centre for Epidemiology, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Oxford Road, Manchester M13 9PT, UK. E-mail: will.dixon@manchester.ac.uk

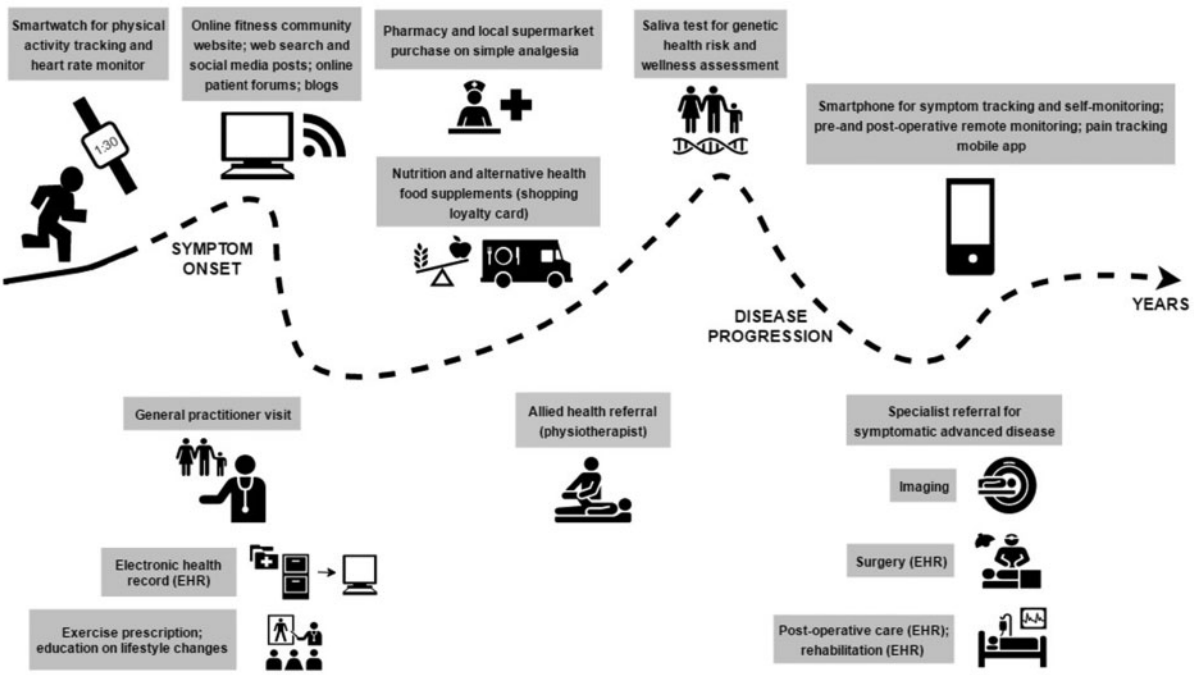
Austin is a middle-aged man who has been previously fit and well, albeit slightly overweight. He enjoys running three times a week and tracks his activity, performance and heart rate using a heart rate monitor linked to his smartwatch, allowing him to understand his achievements and progression. He is in his late 40s, and he has begun to experience persistent knee pain. His family history of OA and discussions on his online fitness community site make him wonder whether he is developing arthritis. He seeks information to learn more about his symptoms through web searches and online forums about OA. He sees his general practitioner (GP) about his knee pain. A diagnosis of knee OA is confirmed and he is referred to a physiotherapist. He is given an exercise prescription, education on lifestyle modification and advice on simple analgesia. He buys over-the-counter (OTC) paracetamol and topical NSAID gel at his local supermarket. Follow-up visits are arranged to assess his knee OA progression and management. In the meantime, he continues to self-monitor and tries to identify possible triggers for days that are worse. He sends his saliva to a commercial company to generate a genetic health risk and wellness report. He experiments with nutritional supplements and alternative health food.

Over the next few years, he notes a progressive deterioration in his performance and an increase in pain. He is referred to an orthopaedic surgeon where he is assessed, has further input from the physiotherapist, has X-rays confirming disease progression and is given a dedicated smartphone application (mobile app) for his pre- and post-operative assessment. He agrees to undergo knee replacement surgery, but the outcome is not what he expected. Despite an initial good recovery, his knee pain persists, which is evident from his pain-tracking mobile app. Given his ongoing frustration with an outcome that did not meet his expectations, he becomes an avid blogger in sharing his post-surgical knee pain experience with others.

Throughout Austin's journey, digital data are captured that could build a picture of his disease and its antecedents, treatments and outcome (Fig. 1). Each discrete data source would show part of his journey but, if linked, could show a more comprehensive picture. These data include what we might consider traditional health data: primary and secondary electronic health records (EHRs) from his GP, specialists and allied health professionals, as well as imaging data. The clinician-derived data are supplemented by patient-generated health data within the health care system, such as his peri-operative mobile app.

There are also recorded data from the patient outside of clinical encounters, including his wearable tracking record, web search history, supermarket loyalty card, genetic profile and social media posts. Collectively these longitudinal types of digital data can provide real-time tracking of symptom trajectory and disease progression and outcome. This opportunity will support individualized and evidence-based understanding of our patients, not only in terms of disease impact on their general well-being, but also of their digitalized information-seeking behavior. Analyses of observational data from some of these sources, such as primary care EHRs, already

Fig. 1 Digital health data capture: a hypothetical case study of the onset and progression of OA



EHR: Electronic health record.

contribute to our understanding of disease and RWE [3–5]. There are many others, like sensors and social media posts, that offer potential RWE, although experience of how to access and analyze such data is, at present, limited.

Observational data in knee OA have indeed added considerably to our knowledge. Age is one of the strongest risk factors for knee OA, perhaps because of an accumulation of other risk factors with time coupled in with the aging process and a reduced ability to withstand adversity in the joint [6]. Obesity is another well-established local risk factor for knee OA [7–9]. Existing knowledge about OA risk factors is often restricted to insights derived from data that are routinely collected, such as age, gender, weight, smoking status, family history and other comorbidities [9]. There is a paucity of evidence about some less readily available risk factors, such as physical activity, diet, other lifestyle factors and health-seeking behavioral, in influencing the development and progression of knee OA. Physical activity is a putative risk factor for knee OA, yet studies provide conflicting evidence [10–12]. This may be due in part to the challenge of accurately measuring and summarizing physical activity patterns over many years before the onset of disease. Prior studies examining this question have used crude measures such as a single self-reported physical activity questionnaire or by comparing, say, runners to non-runners [13, 14]. Yet, as seen in Austin's case study, people are leaving behind them very detailed, digital traces of their active and sedentary living. If these can be linked to health and disease onset and outcomes, we might significantly advance our learning about the risk factors for the onset of knee OA.

Let us now consider Austin's surgery. In Austin's case, why did his total knee replacement (TKR) have a worse outcome compared with similar patients? Many observational studies have examined predictors of poor outcome in TKR. Studies have previously identified factors such as higher pre-operative pain and functional limitations, social disadvantage, depression and anxiety, higher fatigue and higher illness-related distress and co-existing medical conditions [15–17]. Yet Austin had good mental health, high socio-economic status and few comorbidities—so why him? It is possible that he had other important predictors of a poor outcome that may not be easily identified through these study designs. These might include issues such as the timing of surgery with respect to his functional deterioration or his post-operative exercise and other relevant activities during his surgical recovery and rehabilitation. Such metrics and data were collected within his personal digital history but are not yet commonly analyzed across large populations.

As well as contributing to population-level RWE, Austin's use of technology might also support timely interventions. When Austin is faced with intractable pain post-TKR, might his pain be better managed if his day-to-day pain-tracking data, medication use and physical activity data could be accessed? These data might allow his treatment to be personalized instead of escalating the dose and strength of his analgesia and giving generic advice on exercise and self-care. Consumer devices such as activity trackers already employ smart coaching techniques to encourage greater physical activity, guided by contemporaneous data collected on the device. In time, it is possible that real-time analysis of this RWE could lead to personalized digital health interventions, such as post-operative coaching, to support usual clinical care.

Challenges in providing digitalized RWE in future health care and research

New data types

Research that uses the novel data sources described in Austin's case study is still in its infancy. Studies in knee OA are starting to use sensors to evaluate mobility, demonstrating this method of assessment is feasible and may be cost-effective [18–20]. Longitudinal studies using accelerometers are starting to collect much more granular information about physical activity in patients with knee OA, such as the frequency, intensity, time and type of activity [21]. Many such studies provide participants with a research device to track their activity but do not yet provide RWE in free-living individuals without an associated research infrastructure. Although large-scale bespoke research studies using loaned accelerometers can be done, as seen in the UK Biobank study, which provided wrist-worn devices to >100 000 participants, such efforts are a major undertaking [22]. Research using physical activity measured using consumers' own devices is starting to emerge—and sometimes on a large scale. Using data from the Argus app by Azumio (Palo Alto, CA, USA), researchers compared physical activity levels in 717 527 people from 111 countries across the globe [23].

Despite offering big promise, there are open questions such as the validity and quality of the activity measurement and possible selection bias of smartphone and app users. In studies comparing step counts from consumer wearables and smartphones in healthy adults, variability in step count accuracy has been seen between devices [24]. In knee OA specifically, small feasibility studies are exploring whether patterns of physical activity can be collected using raw accelerometer outputs alongside self-reported data using consumer cellular smartwatches [25]. This would make the derivation of a physical activity metric more transparent and standardized and could potentially lead to a future where we are able to have detailed daily information about disease symptoms and progression collected on a single device. The gradual introduction of patient-generated health data from consumer devices into clinical care will lead to significant opportunities for research due to the additional non-clinical context and information available from linkable clinical records data. The overlap between what data and information could, in theory, support both clinical care and research is significant, meaning careful design of systems to meet both needs would deliver multiple benefits [26, 27].

The science of analysing user-generated data from web searches and social media posts is new but has great potential [28]. There are already examples of both promising insight and notable errors from this data mining approach. One of the most highly cited examples of web data analytics was the Google Flu Trends service, which was initially heralded as an exemplary use of big data but was later found to generate inaccurate predictions [29, 30]. Other areas of social media mining for health beyond disease surveillance have included pharmacovigilance and behavioral medicine [31]. Both of these areas have relevance to our case study. Could analysis of OTC NSAID use, captured through store card data or self-reported information, tell us about its efficacy, or could studying paracetamol consumption shine further light on controversies about its effectiveness and safety [32]? Paracetamol safety is notoriously difficult to study using existing data sources such as administrative databases or primary care databases, because they do not capture OTC use that accounts for the vast majority of paracetamol use. As discussed above, how does physical activity influence the onset and outcome of OA? There is an increasing range of research that has explored unstructured data obtained from social media platforms in different rheumatic and musculoskeletal conditions [33–40]. For example, analysis of gout-related social media posts has shown patients are more interested in symptom uncertainties and treatment and less so in serological results of urate and its treat-to-target level [33]. Other studies using social media have brought to light patients' concerns about treatment, for example about biologics or prednisolone therapy [35, 41, 42]. Patients can find it difficult to discuss certain views openly with their clinicians, so analysis of their views captured digitally outside of the clinic consultation can be insightful.

Fundamentals of epidemiology for a digital age: selection bias, validity, missing data and more

The new world of digital health data to support observational research requires us to revisit several fundamentals of epidemiology. Selection bias can be easily introduced, as digital health studies may recruit specific types of participants, such as individuals who are health conscious, digitally literate or have a higher socio-economic class. The validity of new data collection tools needs to be considered, whether it is a digital version of a traditional measure such as a visual analogue scale, a new instrument for self-reported data, an active task such as an app-directed 6-minute walk test or raw or processed passively collected sensor output. For example, can we trust range of motion as measured by a user holding a smartphone and following instructions on a mobile app [43]? Processing and analysis of such data require the establishment of new standards with transparent reporting. For unstructured data, there is a technical challenge in accurately converting these data into structured forms ready for population-level analysis. Researchers may frequently face issues of missing data and diminishing data over time, as a gradual downtrend of user engagement is commonly seen over time in many mobile health studies [43, 44]. Handling continuous streams of sensor data will require a new analysis method not previously required for many epidemiological studies. There have been enormous advances in artificial intelligence (AI) and machine learning in recent years. Progress in non-health industries, such as financial services, the development of driverless cars, speech recognition within smartphones and fraud detection in insurance, has been exceptionally rapid and fruitful. Similar progress is now starting to appear in health care [45, 46]. In OA, the use of AI in automated multidimensional imaging analysis may allow complex computational interpretation and aggregation of these sophisticated imaging data, linking them to patient-generated health data and clinical care data [47]. There are concerns that predictive analytics in health, despite good model performance, may not be sufficiently transparent to enable clinical buy-in and trust, a challenge that may be helped by emerging developments in 'interpretable' machine learning or 'explainable AI' [48].

Governance and public trust for real-world digital health data

Each discrete data source described above and its associated analysis is promising, yet it is clear that the real value will come when different digital data sources can be collated to give a more comprehensive picture.

Appropriate governance on data ownership and data protection is imperative as we move towards the idea of acquiring, aggregating and archiving linked digital health data. Patients should be able to control their own data, with clarity about who has handled their data. This will allow them freedom and rights in protecting, linking and sharing their data with other digital health users, such as health care professionals and researchers. It is inevitable that only certain members of society would wish to have their data shared in this way for research, and their views should be heard. Yet this should not necessarily prevent any data sharing within society. Initiatives like CitizenMe allows individuals to store their digital data in their personal data cloud, as well as to participate in surveys which may provide small monetary incentives [49].

Official regulation in reinforcing data security, consent for data linkage and privacy is important in the digital era [50]. Standards and guidelines are emerging, but it remains a gray area at times. In the UK, the use of anonymized EHR data for research does not require patient consent [51]. In primary care research databases such as the UK's Clinical Practice Research Datalink, the data controller does not hold identifiable patient information and therefore cannot facilitate contact with patients. Nonetheless, it is possible for studies to collect data from patients alongside their EHR data via their GP, albeit with limited uptake from practices and patients [52].

Keeping public trust can sometimes require more than abiding by governance regulations, and so researchers must be thoughtful about how they clearly communicate the benefits and managed risks of data sharing [53]. Analysis of social media data still requires care, as users

may not understand that their data are publically available and may not wish their data to be used for research [54]. Future digital health and social care data require a bona fide and secure infrastructure for data storage and use. As outlined by Mandl and Kohane [55, 56], standardization and interoperability of different digital data sources are crucial for ensuring correct and valid data acquisition from patients and appropriate implementation of these data in self-care, clinical care and research. When collecting consent on digital devices, a new model of consent is required in the absence of study nurses, traditional consent forms and patient information sheets. Guidelines for electronic consent have been published by the US Food and Drug Administration, and there are already examples of excellent practice in mobile health studies such as the MyHeart Counts cardiovascular health study [44, 57].

While there remain lots of challenges in the area of governance, citizen consent and privacy are well delivered in many other aspects of our digital lives, such as banking, so strong governance of health data with public trust should be entirely possible. Recent initiatives such as the Wellcome Trust's Understanding Patient Data, Health Data Research UK and the new Ada Lovelace Institute, an independent research and deliberative body with a mission to ensure data and AI work for people and society, will help ensure public trust remains at the forefront of developments in health data and research [53, 58]. As stated in the report from the Select Committee on Artificial Intelligence, 'maintaining public trust over the safe and secure use of their data is paramount to the successful widespread deployment of AI and there is no better exemplar of this than personal health data' [59].

Conclusion

In summary, data about the causes and determinants of disease and its outcome are increasingly being collected digitally. It is already possible to see that such data will be hugely valuable. We are moving from a time when disease could be measured only at sparse intervals, such as at a 6-month clinic appointment, to a situation where many aspects and correlates of disease can be tracked frequently or for the first time. Novel data types provide an opportunity to answer questions that were previously difficult or impossible to answer. Yet there remain significant challenges around the appropriate governance of such data that maintains public trust and how we ensure we derive appropriate insight given the representativeness of the new digital patient. The inevitable move into the digital era means we should embrace, rather than hide behind, these challenges and ensure we make the best use of the opportunities that this new RWE presents to us.

Acknowledgements

H.L.P. is the recipient of the 2018 Arthritis Australia National Research Program – Ken Muirden Overseas Training Fellowship, an educational research grant funded by the Australian Rheumatology Association.

Funding: This work was supported by the Arthritis Research UK Centre for Epidemiology (21755).

Disclosure statement: W.G.D. has received consultancy fees from Bayer and Google. H.L.P. declared no conflicts of interest.

References

1 Gartner. Gartner says solving ''Big Data'' challenge involves more than just managing volumes of data. 2011. https://www.gartner.com/newsroom/id/1731916 (25 September 2018, date last accessed).

2 Schroeck M, Shockley R, Smart J, Romero-Morales D, Tufano P. Analytics: the real-world use of big data. IBM Institute for Business Value. 2012. https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=GBE03519USEN (25 September 2018, date last accessed).

3 Arden N, Altman D, Beard D *et al*. Lower limb arthroplasty: can we produce a tool to predict outcome and failure, and is it cost-effective? An epidemiological study. Programme Grants for Applied Research. Southampton (UK): NIHR Journals Library, 2017.

4 Culliford D, Maskell J, Judge A *et al*. Future projections of total hip and knee arthroplasty in the UK: results from the UK Clinical Practice Research Datalink. Osteoarthritis Cartilage 2015;23:594–600.

5 Khan T, Alvand A, Prieto-Alhambra D *et al*. ACL and meniscal injuries increase the risk of primary total knee replacement for osteoarthritis: a matched case-control study using the Clinical Practice Research Datalink (CPRD). Br J Sports Med 2018. doi: 10.1136/bjsports-2017-097762.

6 Busija L, Bridgett L, Williams SR *et al*. Osteoarthritis. Best Pract Res Clin Rheumatol 2010;24:757–68.

7 Cooper C, Snow S, McAlindon TE *et al*. Risk factors for the incidence and progression of radiographic knee osteoarthritis. Arthritis Rheum 2000;43:995–1000.

8 Reijman M, Pols HA, Bergink AP *et al*. Body mass index associated with onset and progression of osteoarthritis of the knee but not of the hip: the Rotterdam Study. Ann Rheum Dis 2007;66:158–62.

9 Silverwood V, Blagojevic-Bucknall M, Jinks C *et al*. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. Osteoarthritis Cartilage 2015;23:507–15.

10 Richmond SA, Fukuchi RK, Ezzat A *et al*. Are joint injury, sport activity, physical activity, obesity, or occupational activities predictors for osteoarthritis? A systematic review. J Orthop Sports Phys Ther 2013;43:515–B19.

11 Tran G, Smith TO, Grice A *et al*. Does sports participation (including level of performance and previous injury) increase risk of osteoarthritis? A systematic review and meta-analysis. Br J Sports Med 2016;50:1459–66.

12 Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Clin Geriatr Med 2010;26:355–69.

13 Chakravarty EF, Hubert HB, Lingala VB, Zatarain E, Fries JF. Long distance running and knee osteoarthritis. A prospective study. Am J Prev Med 2008;35:133–8.

14 McAlindon TE, Wilson PW, Aliabadi P, Weissman B, Felson DT. Level of physical activity and the risk of radiographic and symptomatic knee osteoarthritis in the elderly: the Framingham Study. Am J Med 1999;106:151–7.

15 Judge A, Arden NK, Cooper C *et al*. Predictors of outcomes of total knee replacement surgery. Rheumatology (Oxford) 2012;51:1804–13.

16 Lindberg MF, Rustoen T, Miaskowski C, Rosseland LA, Lerdal A. The relationship between pain with walking and self-rated health 12 months following total knee arthroplasty: a longitudinal study. BMC Musculoskelet Disord 2017;18:75.

17 Lingard EA, Katz JN, Wright EA, Sledge CB. Predicting the outcome of total knee arthroplasty. J Bone Joint Surg Am 2004;86:2179–86.

18 Calliess T, Bocklage R, Karkosch R *et al*. Clinical evaluation of a mobile sensor-based gait analysis method for outcome measurement after knee arthroplasty. Sensors (Basel) 2014;14:15953–64.

19 Staab W, Hottowitz R, Sohns C *et al*. Accelerometer and gyroscope based gait analysis using spectral analysis of patients with osteoarthritis of the knee. J Phys Ther Sci 2014;26:997–1002.

20 Verlaan L, Bolink SA, Van Laarhoven SN *et al*. Accelerometer-based physical activity monitoring in patients with knee osteoarthritis: objective and ambulatory assessment of actual physical activity during daily life circumstances. Open Biomed Eng J 2015;9:157–63.

21 Sliepen M, Mauricio E, Lipperts M, Grimm B, Rosenbaum D. Objective assessment of physical activity and sedentary behaviour in knee osteoarthritis patients - beyond daily steps and total sedentary time. BMC Musculoskelet Disord 2018;19:64.

22 Doherty A, Jackson D, Hammerla N *et al*. Large scale population assessment of physical activity using wrist worn accelerometers: the UK Biobank study. PLoS One 2017;12:e0169649.

23 Althoff T, Sosic R, Hicks JL *et al*. Large-scale physical activity data reveal worldwide activity inequality. Nature 2017;547:336–9.

24 Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA 2015;313:625–6.

25 Beukenhorst AL, Sergeant JC, Little MA, McBeth J, Dixon WG. Consumer smartwatches for collecting self-report and sensor data: app design and engagement. Stud Health Technol Inform 2018;247:291–5.

26 Dixon WG, Michaud K. Using technology to support clinical care and research in rheumatoid arthritis. Curr Opin Rheumatol 2018;30:276–81.

27 Wu AW, Kharrazi H, Boulware LE, Snyder CF. Measure once, cut twice – adding patient-reported outcome measures to the electronic health record for comparative effectiveness research. J Clin Epidemiol 2013;66(8 Suppl):S12–20.

28 Estrin D. Small data, where n = me. Commun ACM 2014;57:32–4.

29 Ginsberg J, Mohebbi MH, Patel RS *et al*. Detecting influenza epidemics using search engine query data. Nature 2009;457:1012–4.

30 Lazer D, Kennedy R, King G, Vespignani A. The parable of Google Flu: traps in big data analysis. Science 2014;343:1203–5.

31 Paul MJ, Sarker A, Brownstein JS *et al*. Social media mining for public health monitoring and surveillance. Pacific Symp Biocomput 2016;468–79.

32 Roberts E, Delgado Nunes V, Buckner S *et al*. Paracetamol: not as safe as we thought? A systematic literature review of observational studies. Ann Rheum Dis 2016;75:552–9.

33 Derksen C, Serlachius A, Petrie KJ, Dalbeth N. ''What say ye gout experts?'' a content analysis of questions about gout posted on the social news website Reddit. BMC Musculoskelet Disord 2017;18:488.

34 des Bordes JKA, Gonzalez E, Lopez-Olivo MA *et al*. Assessing information needs and use of online resources for disease self-management in patients with rheumatoid arthritis: a qualitative study. Clin Rheumatol 2018;37:1791–7.

35 Dzubur E, Khalil C, Almario CV *et al*. Patients' concerns and perceptions regarding biologic therapies in ankylosing spondylitis: insights from a large-scale survey of social media platforms. Arthritis Care Res (Hoboken) 2019;71:323–30.

36 Geuens J, Swinnen TW, Westhovens R *et al*. A review of persuasive principles in mobile apps for chronic arthritis patients: opportunities for improvement. JMIR mHealth uHealth 2016;4:e118.

37 Meldrum S, Savarimuthu BT, Licorish S *et al*. Is knee pain information on YouTube videos perceived to be helpful? An analysis of user comments and implications for dissemination on social media. Digit Health 2017;3: 1–18.

38 Merolli M, Gray K, Martin-Sanchez F, Lopez-Campos G. Patient-reported outcomes and therapeutic affordances of social media: findings from a global online survey of people with chronic pain. J Med Internet Res 2015;17:e20

39 Modica RF, Lomax KG, Batzel P *et al*. The family journey-to-diagnosis with systemic juvenile idiopathic arthritis: a cross-sectional study of the changing social media presence. Open Access Rheumatol 2016;8:61–71.

40 Sunkureddi P, Doogan S, Heid J *et al*. Evaluation of self-reported patient experiences: insights from digital patient communities in psoriatic arthritis. J Rheumatol 2018;45:638–47.

41 Costello R, Patel R, Humphreys J, McBeth J, Dixon WG. Patient perceptions of glucocorticoid side effects: a cross-sectional survey of users in an online health community. BMJ Open 2017;7:e014603.

42 Patel R, Belousov M, Jani M *et al*. Frequent discussion of insomnia and weight gain with glucocorticoid therapy: an analysis of Twitter posts. npj Digit Med 2018;1:1–7.

43 Crouthamel M, Quattrocchi E, Watts S *et al*. Using a ResearchKit smartphone app to collect rheumatoid arthritis symptoms from real-world participants: feasibility study. JMIR mHealth uHealth 2018;6:e177.

44 McConnell MV, Shcherbina A, Pavlovic A *et al*. Feasibility of obtaining measures of lifestyle from a smartphone app: the MyHeart Counts cardiovascular health study. JAMA Cardiol 2017;2:67–76.

45 De Fauw J, Ledsam JR, Romera-Paredes B *et al*. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med 2018;24:1342–50.

46 Rajkomar A, Oren E, Chen K *et al*. Scalable and accurate deep learning with electronic health records. npj Digit Med 2018;1:1–10.

47 Pedoia V, Majumdar S. Translation of morphological and functional musculoskeletal imaging. J Orthop Res 2018;37:23–34.

48 Pavlus J. A new approach to understanding how machines think. Quanta Magazine 10 January 2019. https://www.quantamagazine.org/been-kim-is-building-a-translator-for-artificial-intelligence-20190110/# (13 January 2019, date last accessed).

49 CitizenMe. Get real life value from your life online. 2018. https://www.citizenme.com/public/wp/ (19 September 2018, date last accessed).

50 Chan T, Di Iorio CT, De Lusignan S *et al*. UK national data guardian for health and care's review of data security: trust, better security and opt-outs. J Innov Health Inform 2016;23:627–32.

51 National Data Guardian for Health and Care. Review of data security, consent and opt-outs. 2016. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/535024/data-security-review.PDF (19 September 2018, date last accessed).

52 Joseph RM, Soames J, Wright M *et al*. Supplementing electronic health records through sample collection and patient diaries: a study set within a primary care research database. Pharmacoepidemiol Drug Saf 2018;27:239–42.

53 Wellcome Trust. Understanding patient data. https://understandingpatientdata.org.uk (25 September 2018, date last accessed).

54 Hudson JM, Bruckman A. ''Go away'': participant objections to being studied and the ethics of chatroom research. Inform Soc 2004;20:127–39.

55 Mandl KD, Kohane IS. Time for a patient-driven health information economy? N Engl J Med 2016;374:205–8.

56 Mandl KD, Kohane IS. A 21st-century health IT system—creating a real-world information economy. N Engl J Med 2017;376:1905–7.

57 US Department of Health and Human Services. Use of electronic informed consent: questions and answers. Guidance for institutional review boards, investigators, and sponsors. Washington, DC: US Department of Health and Human Services, 2016.

58 Nuffield Foundation. Ada Lovelace Institute. https://www.adalovelaceinstitute.org/ (13 January 2019, date last accessed).

59 Select Committee on Artificial Intelligence. AI in the UK: ready, willing and able? HL Paper 100. London: House of Lords. 2018. https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf (13 January 2019, date last accessed).

3.4 Summary

This chapter summarises the ever-expanding volume and breadth of digitalised health-related data in the provision of real-world evidence in clinical practice and research, with the use of a clinical vignette of a patient with OA for the topic discussion.

Main findings include:

- New types of health-generated data (structured or unstructured) obtained digitally through consumer's own device offer big promise in research, real-time intervention, and self-management
- Information obtained from digitalised health-related data may raise questions on the validity and quality of the data collected, and empowers the development of novel methods in data transformation and analysis when evaluating unstructured data
- Fundamentals of epidemiology and transparency are instrumental in the processing and analysis of such data
- Reinforcement in data security, consent, and privacy is imperative in the digital era, and governance in data protection and ownership is necessary as these data are acquired, aggregated, and archived.

Chapter 4: The Application of Intensive Longitudinal Methods in Analysing Pain Using Mobile Health Data

4.1 Preface

This chapter presents the research work and skills I achieved during the first 15 months of my candidature in Manchester, UK. In early 2018, I had the opportunity to be involved in one of the largest mobile health studies in the UK, *Cloudy with a Chance of Pain*, led by Professor William Dixon (my external supervisor) and his team. *Cloudy with a Chance of Pain* was the first smartphone study conducted to examine the association of weather and pain.

During this candidature period, I developed my foundation in epidemiology knowledge and research skills, in particular statistical programming skills in R, Stata, and Python. In addition, I had the privilege to work with a multidisciplinary team across UK and internationally, which included epidemiologists, rheumatologists, meteorologists, statisticians, mathematicians, postgraduate students, project manager, and the patient and public engagement representatives.

In this chapter, I first present the background for examining longitudinal data on pain symptoms and the analysis of pain trajectory and daily pain variability using intensive longitudinal methods, thus forming the primary aims of this exploratory work. I present the overview of the study, Cloudy with a Chance of Pain, in detail, as well as the subset of data that I used for my study analysis. The methodology and results of my exploratory analyses on pain trajectory and daily pain variability of individuals with inflammatory and non-inflammatory RMDs are presented. Next, an overall synthesis and discussion of the study results are presented, with emphasis on the methods used and the epidemiological perspectives. This work is prepared and written in a conventional thesis chapter format.

Research skills achieved during this candidature period is discussed in detail in Chapter 8. Statistical code scripts for my study analyses are included in Appendix C. Additionally, as I was involved as a co-author in the primary publications for this study, the published articles are included in Appendix D.

4.2 Background

An important aspect of understanding the variability of pain experienced, often not well captured in traditional pain trajectory analysis, is the extent and frequency of pain fluctuation over time. Such understanding of pain variability is essential to gather information of spontaneous pain moments as seen in flares in RA, OA, or LBP. Episodes of flare are often the impetus for guiding treatment decision and pain management in a typical rheumatology clinic consultation.

Individuals living with chronic pain from their RMDs often describe a spectrum of pain intensity, duration, and frequency, depending on other co-occurring symptoms and activities related to the pain experience. In such phenomenology of chronic pain, its temporal variability, especially capturing those spontaneous or unprovoked moments of pain with varying duration and frequency, either good or bad, remains an ongoing topic of interest for many in understanding the dynamics of pain across different RMDs.

Broadly, chronic pain in RMDs is dynamic and varies at different timepoint of measurements. Often, the longitudinal measure of these changes of pain over time, known as pain trajectory, is studied across different RMDs. Statistically, with the assumption under a normal distribution, at least three repeated measures of pain are adequate to define an individual's pain trajectory over a pre-defined period[226]. Population-averaged or aggregated summary course of pain is often insufficient to describe the temporal dynamics of the pain severity, as described in reviews summarising the work done in studies investigating the course of LBP and pain trajectory in knee OA[152, 227]. Fundamentally, examining the temporal dynamics of pain often requires the evaluation of the variability of pain, the context, the diurnal pattern, and the activities corresponding to the pain level measured in real-time[178, 179, 197, 198, 228-232]. In LBP, different pain trajectory phenotypes were observed, with common pain patterns identified across different studies, regardless of settings, countries, and frequency of pain measures[152, 233-242]. Similarly, studies examining

pain trajectories in early to moderate knee OA had identified common pain trajectory phenotypes, with a stable pain trajectory found in 85% of the aggregated study population, and relevant predictors of worsening pain seen such as depressed mood, low education, and higher comorbidities[205, 227, 243-248]. Common statistical methods in clustering analysis such as latent class analysis, latent class growth analysis and hierarchical class analysis were commonly used in these studies. Such data-driven analytical approach allows distinct stratification of pain trajectory phenotypes, with a birds-eye view in differentiating at-risk individuals with unstable or worsening pain trajectories and individuals with stable or improving pain trajectories (Figure 4.1).

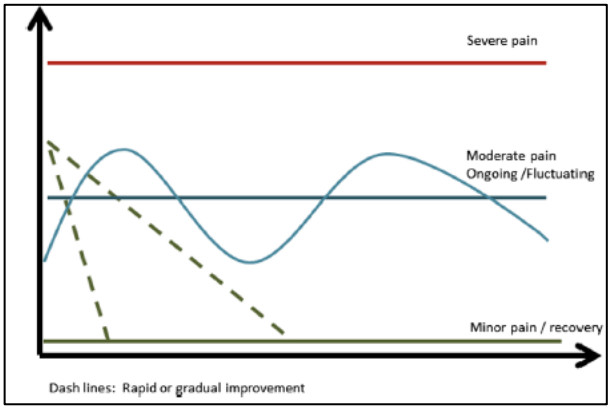


Figure 4.1 A simplified diagram of pain trajectory phenotypes in low back pain

Source: Adapted from Kongsted A, Kent P, Axen I, et al[152] (used with permission under Creative Commons Attribution License)

As shown in Figure 4.2, episodic or fluctuating pain patterns in LBP may differ in pain intensity and frequency of change in pain between individuals, and yet, these individuals were grouped in the same pain trajectory phenotype[152]. Examining the intraindividual differences in the dynamic of pain in RMDs provides a detailed snapshot of the components of pain variability, in terms of assessing the extent and frequency of pain fluctuations across time and contextual settings[219].

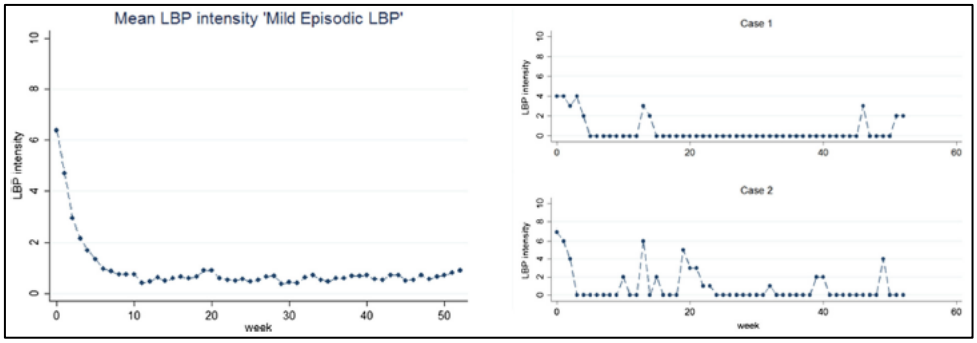


Figure 4.2 An example of mild episodic pain trajectory phenotype in low back pain (left), with individual cases (right) of mildly episodic pain trajectory in Case 1 and highly episodic pain trajectory in Case 2

Source: Adapted from Kongsted A, Kent P, Axen I, et al[152] (used with permission under Creative Commons Attribution License)

Real-time data-driven analysis of change in pain intensity over time captured through EMA, as discussed in Chapter 1, had identified relevant components of pain variability using time series-based regime switching model[198]. In this model, Schneider and colleagues stated these identified temporal dynamic measures of pain as 'Average pain (mean level over time), Amplitude (magnitude of shifts in pain levels), Persistence (average duration of pain states), and Dominance (relative duration of higher vs lower pain states)', as seen in Figure 4.3[198].

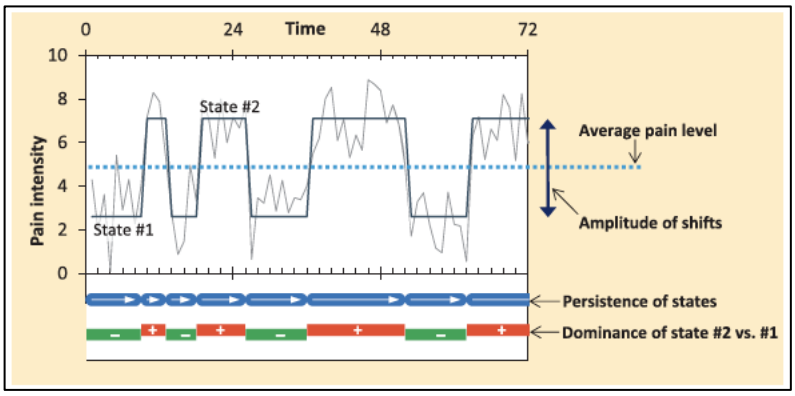


Figure 4.3 Measures of temporal dynamics of pain using regime-switching model

Source: Adapted from Schneider S, Junghaenel DU, Ono M, et al[198] (used with permission under the Wolters Kluwer Health, Inc – Copyright Clearance Center's Rightslink)

As described earlier in Chapter 3, in the presence of any chronic condition, digitalised patient-generated health data can provide invaluable information on many aspects of the individual's well-being state as well as real-time capture of disease symptoms and disease impact. For example, in a typical rheumatology clinic, will the clinician have a better understanding of a patient's pain experience outside of the clinic setting, if such real-time data is captured using one of the mHealth devices? Additionally, once the data are available for an in-depth analysis, how can we best assess such intensive data appropriately, and yet, relevant to the patient and the treating clinician? Currently, little is known on the applicability of such intensive longitudinal methods in examining change in pain over time captured through real-time mHealth app. Therefore, in this chapter, by using the daily data on pain symptoms captured in *Cloudy with a Chance of Pain*, I present my research work in analysing daily pain trajectory and pain variability across different RMDs.

4.3 Study Aims and Study Outline

In this exploratory work using data from *Cloudy with a Chance of Pain*, a national UK smartphone study, and using intensive longitudinal methods, I aimed to examine the long-term and short-term day-to-day pain variability in individuals with different types of RMDs (both inflammatory and non-inflammatory).

In detail, for the long-term and short-term pain variability analyses, multilevel model and Markov transition model were applied respectively. I hypothesised that there were differences in the temporal patterns of pain variability among individuals across different RMDs.

I was involved in the *Cloudy with a Chance of Pain* study at the commencement of the primary study analysis of examining the weather-pain relationship, which occurred following the completion of the study design, study app, and study feasibility. In the next subchapter,

the main study dataset of the *Cloudy with a Chance of Pain* is presented, which serves a purpose for the reader to understand the concept of this smartphone study and to further guide the discussion of the methods and analyses of my study in this thesis.

4.4 Dataset: *Cloudy with a Chance of Pain*

Cloudy with a Chance of Pain was an mHealth study designed and conducted in the UK between 2016 and 2017. This large, prospective citizen-scientist cohort study was the largest mHealth national study in the UK to evaluate the weather-pain relationship. Below is an illustrated flow diagram of the study eligibility and recruitment as well as types of data collected in this study (Figure 4.4).

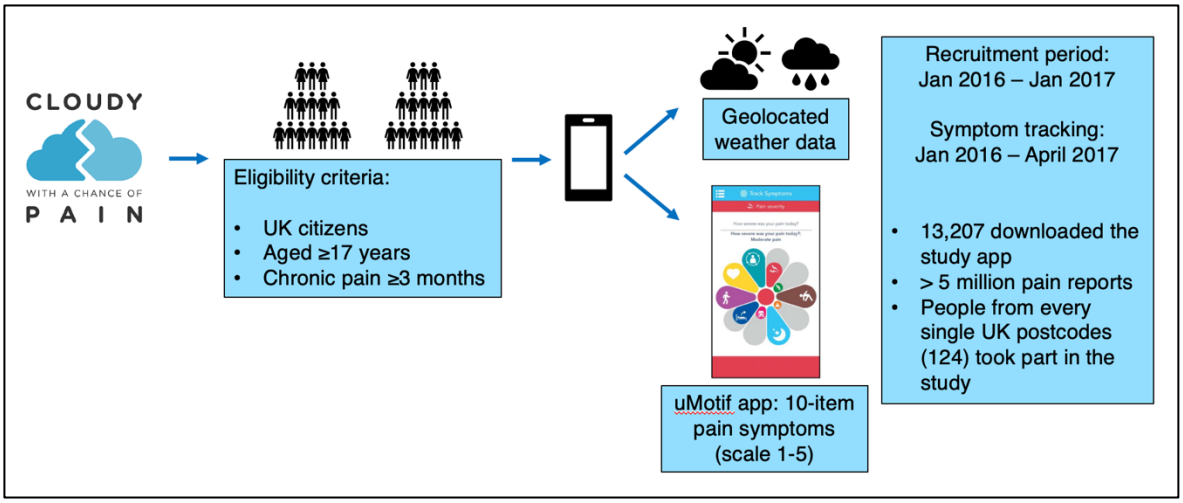


Figure 4.4 A flow diagram summarising the study eligibility and recruitment and the pain and weather data collection

4.4.1 Study recruitment and patient and public involvement

Study recruitment commenced in January 2016 and ended in January 2017, with the opportunity for study participants to continue to enter their daily data on pain symptoms beyond the study recruitment period up until April 2017. At the beginning, the study information reached out to the public through the local and national media broadcasts via

television, radio, and press, as well as social media. Interested smartphone users were guided to the study website www.cloudywithachanceofpain.com, of which the study design, eligibility, study app download and enrolment process were outlined in detail.

Patient and public involvement (PPI) was instrumental to the success of this study prior to the data collection from study participants in the *Cloudy with A Chance of Pain*[249-251]. In detail, the PPI group was involved in the local UK media broadcast highlighting the importance of the study question around weather-pain relationship[249]. The PPI group also took part in co-designing the study app during the feasibility study phase[251]. In addition, the PPI group contributed to the interpretation and dissemination of the study findings to the study participants and the community in general.

4.4.2 Study eligibility and type of data collection

The study eligibility criteria included individuals who were:

- Aged ≥17 years,
- UK residents,
- Having physician-diagnosed chronic pain condition with a duration of at least 3 months, and
- Smartphone users (either iPhone Operating System, iOS or Android)

The 'Cloudy' app was designed and reconfigured from a pre-existing mHealth app by uMotif, a UK-based company focusing on co-designing user-engaging platform for clinical and research in collecting digitalised real-world data[252]. In addition, uMotif was commissioned to the develop the GPS tool, allowing concurrent weather-pain data to be collected. Once downloaded, the app would guide the individual through an informed consent process electronically. Upon registration, study participants were asked to complete the baseline questionnaires (Figure 4.5).

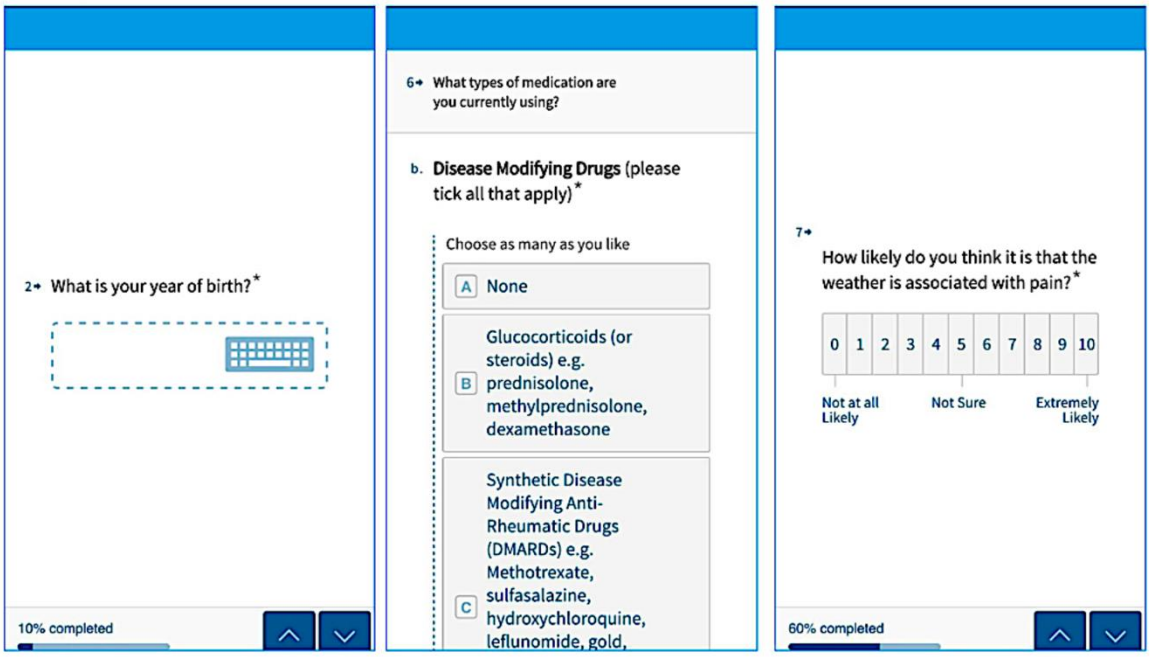


Figure 4.5 A screenshot of some of the baseline questionnaires presented on the study app interface

Source: Adapted from Druce KL, McBeth J, van der Veer SN, et al[253] (used with permission under Creative Commons Attribution License)

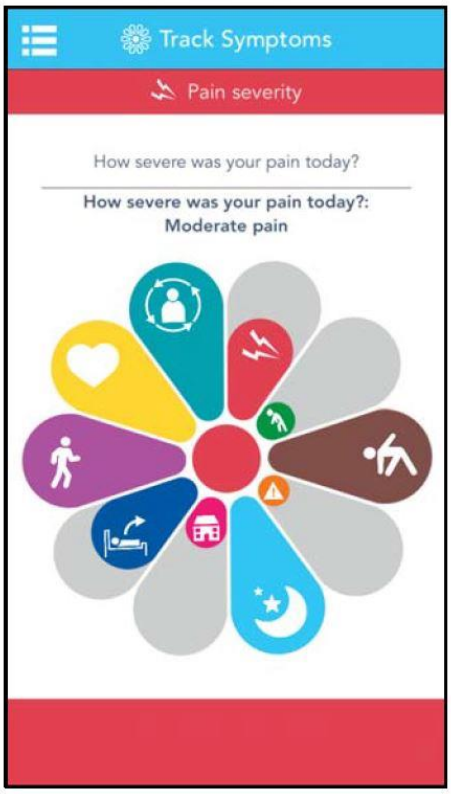
These baseline questionnaires included information on age, gender, the outward code of their residential UK postcode, pain site, clinician-diagnosed pain-related comorbidities, medication use and personal beliefs on weather-pain relationship (Table 4.1).

Following the completion of the baseline questionnaires, study participants were able to access the daily recording of their pain symptoms using an interactive interface on the app. In this daily tracking app, each of the ten self-reported questionnaires on pain symptoms was presented as the 'petal' of a flower-like motif, designed by uMotif (Figure 4.6).

Baseline data information	Question	Response (± options)
Gender	Are you male or female?	• Female • Male
Year of birth	What is your year of birth?	Free text
Residential UK postcode		Free text (outward code only)
Pain site(s)	Where is your pain?	• Mouth or jaw • Neck or shoulder • Back pain • Stomach or abdominal • Hip pain • Knee pain • Hands • Feet • Pain at multiple sites • Pain all over body
Chronic pain condition(s)	Has your doctor ever told you that you have any of the following pain-related conditions? (please tick all that apply)	• No pain • Rheumatoid arthritis • Osteoarthritis • Spondyloarthropathy/ankylosing spondylitis • Gout or other crystal arthritis (e.g., pseudogout) • Arthritis (type not specified) • Chronic headache • Neuropathic pain • Other

Medication(s)	What types of medication are you currently using? Painkillers (please tick all that apply)	• None • Paracetamol • NSAIDs (tablets or creams/ gels) e.g., ibuprofen, diclofenac, naproxen, indomethacin, meloxicam, ketoprofen, celecoxib, etoricoxib, etodolac • Other simple analgesics e.g., cocodamol, codydramol • Weak opiates e.g., codeine, dihydrocodeine, tramadol • Strong opiates e.g., buprenorphine, fentanyl, morphine, oxycodone • Drugs for neuropathic pain e.g., gabapentin, pregabalin, duloxetine • Other
	Disease modifying drugs (please tick all that apply)	• None • Glucocorticoids (or steroids) e.g., prednisolone, methylprednisolone, dexamethasone • Synthetic disease modifying anti-rheumatic drugs (DMARDs) e.g., methotrexate, sulfasalazine, hydroxychloroquine, leflunomide, gold, azathioprine, mycophenolate, cyclosporin, cyclophosphamide • Biologic DMARDs e.g., etanercept (Enbrel), infliximab (Remicade), adalimumab (Humira), certolizumab (Cimzia), golimumab (Simponi), rituximab (Mabthera), tocilizumab (Actemra), abatacept (Orencia) • Other
Personal belief on weather-pain relationship	How likely do you think it is that the weather is associated with pain?	0-10 scale with 0 = not at all likely, 5 = not sure, 10 = extremely likely

Table 4.1 Baseline questionnaires presented on the study app interface



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 4.6 A screenshot of the data item tracking presented on the study app interface (left) and pain symptom reporting (right)

Source: Adapted from Dixon WG, Beukenhorst AL, Yimer BB, et al[250] (used with permission under Creative Commons Attribution License)

In detail, as shown in Figure 4.6, each coloured petal on the interface (left) represented each of the data item (right) and pain symptom was tracked based on participant's response on the interface by dragging the 'petal' from the centre outwards. For each data item, there are five-level ordinal categories of severity, with each level of severity anchoring from bottom to top (from proximal to distal to the centre of the interface), as shown on Figure 4.6. This interactive platform was presented to the study participants daily at 6:24p.m. UK time through a gentle reminder using the 'push notification'.

Alongside real-time capture on daily pain symptoms, concurrent geolocation of weather data was obtained. This was achieved using the GPS embedded within the smartphone as a proxy to link the weather data derived from the nearest weather stations as provided by

the Met Office, a nationwide meteorological service provider in the UK. Data linkage between the weather observation from these weather stations and the weather data from the Integrated Surface Database (ISD) of the National Oceanic and Atmospheric Association (NOAA) was performed. These weather data were important for the primary study analyses of the weather-pain relationship[250, 254, 255].

4.4.3 Study engagement, retention, and cohort selection

During the 12-month study recruitment, across 124 UK postcode areas, a total of 13,207 UK citizens downloaded the readily available 'Cloudy' app using smartphones of either iOS or Android. Approximately 5.1 million longitudinal pain symptoms were captured over this 15-month study period (January 2016 – April 2017).

In the preparation for a final analysis-ready dataset, study participants with complete data on baseline questionnaires and at least one pain data entry were selected. Therefore, 10,584 of the total 13,207 'Cloudy' app users were selected as the study cohort for analysis. Study engagement and retention was described in Figures 4.7 and 4.8.

In addition, the study engagement was initially analysed in a cluster analysis for the first 6-month recruitment period (January-July 2016), as described in detail elsewhere[253]. The study engagement was relatively high, with 6,850 (65%) study participants remained in the study for the first month, and with 4,692 (44%) beyond the first month (not shown). In this cluster analysis using first-order hidden Markov models, within the range of median days of entry of 1-149, four clusters were observed – high (13.60%), moderate (21.76%), low (39.35%), and tourists (25.44%) (Figure 4.8)[250].

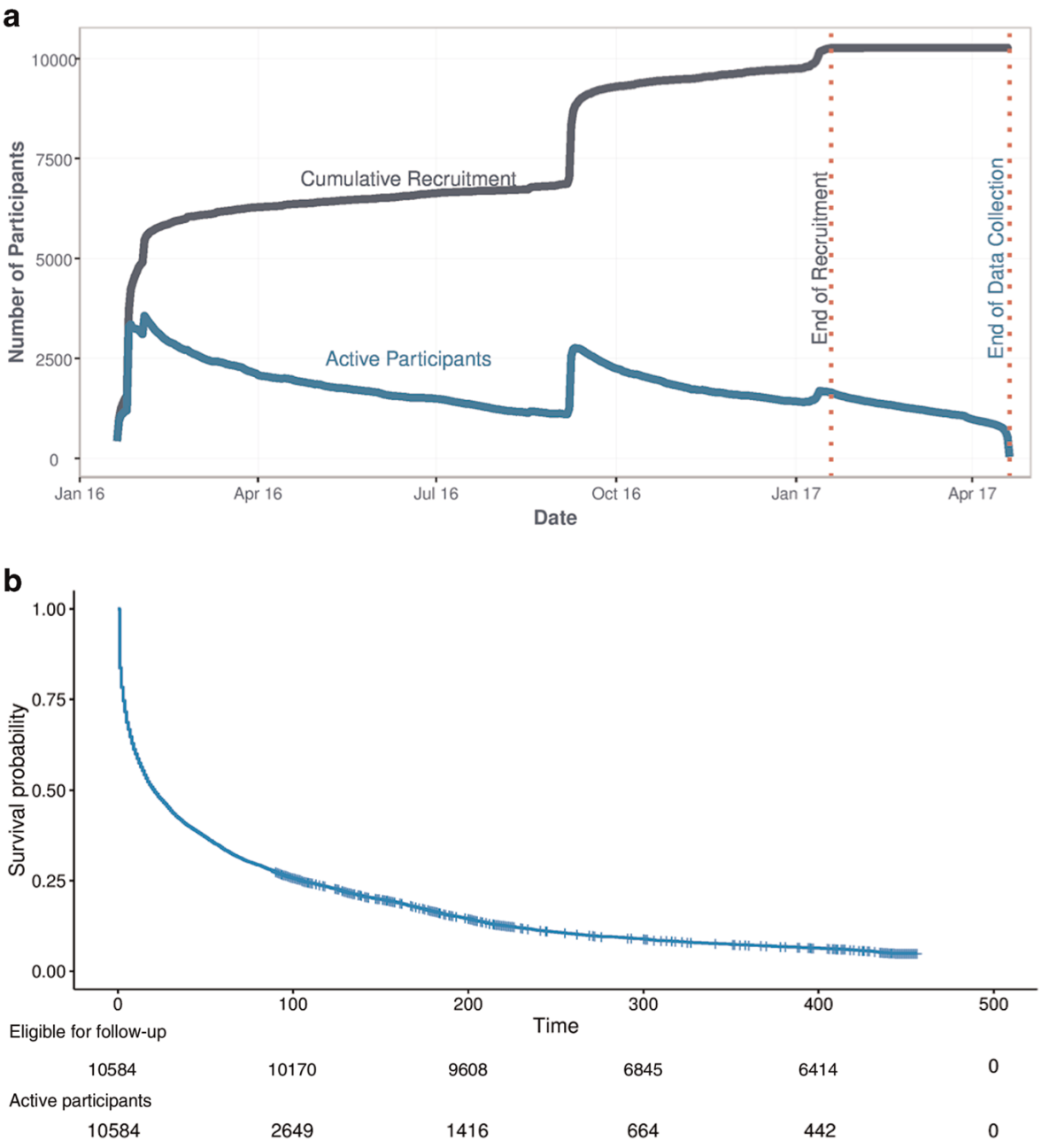


Figure 4.7 Study recruitment (top) and study retention (bottom) throughout the study period. The top figure (a) presented the cumulative recruitment (in black) and active participants remained in the period of study data collection (in blue), with the study end period highlighted for both recruitment (January 2017) and data collection (April 2017) respectively (in red dotted line). In bottom figure (b), using survival probability, active study participants retained in the study over time were outlined (in blue)

Source: Adapted from Dixon WG, Beukenhorst AL, Yimer BB, et al[250] (used with permission under Creative Commons Attribution License)

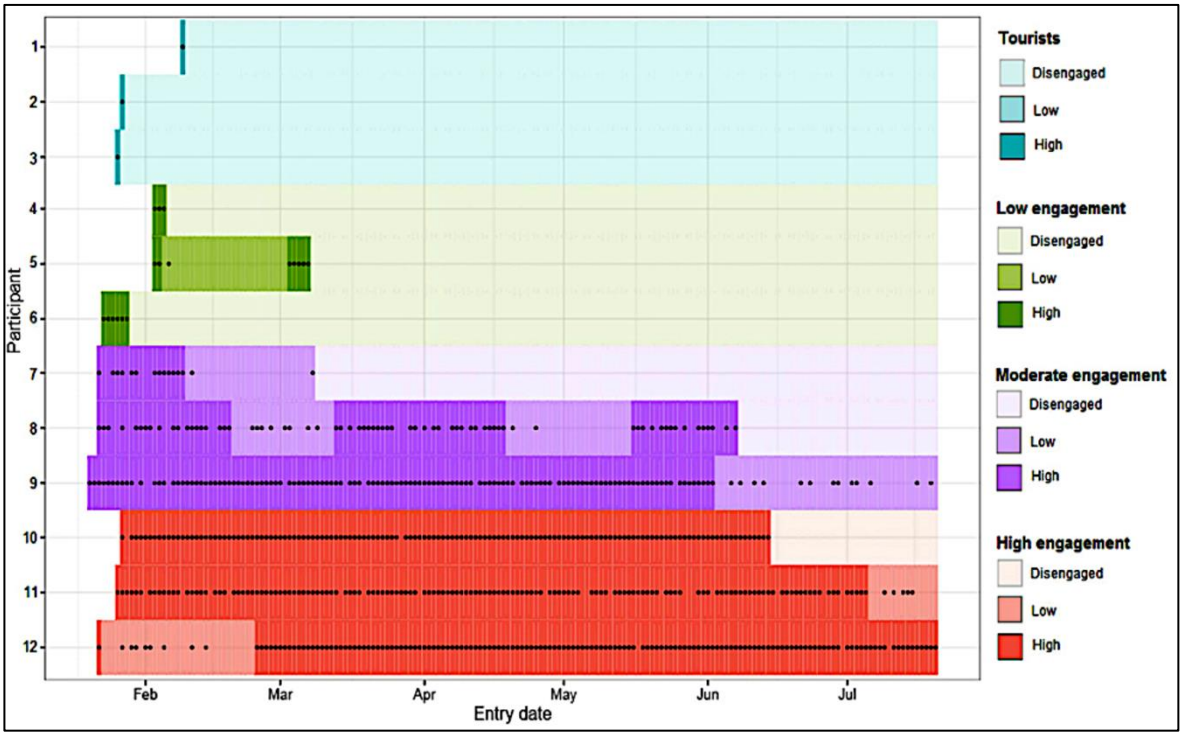


Figure 4.8 Selected few study participants representing each cluster – tourists (in teal), low engagement (in green), moderate engagement (in purple), high engagement (in red)

Source: Adapted from Druce KL, McBeth J, van der Veer SN, et al[253] (used with permission under Creative Commons Attribution License)

4.5 Ethics Approval and Data Management

Ethics approval for this study was obtained from the University of Manchester (Manchester Research Ethics Committee – ethics/15522) and the National Health Service (NHS) Integrated Research Application System (IRAS), UK – 23/NW/0716. The intellectual property agreement for this study remained with the research lead, Professor William Dixon, and the University of Manchester.

In terms of the data management, the study data remained within the secured Information Technology (IT) environment provided by the University of Manchester IT services. I performed my study analyses both within the computer provided by the University of Manchester during my remote candidature period in Manchester, UK, and remotely via a secured Virtual Private Network (VPN)-based virtual computer with statistical support. No

data was removed or used in any form without permission during my research work for this thesis, except for exporting the statistical code scripts and study results specific to my study analyses.

4.6 Methods

4.6.1 Cohort and study period selection

As discussed in Chapter 1, the natural history of chronic pain in patients with inflammatory RMD is mechanistically different from those with non-inflammatory RMD such as fibromyalgia. Furthermore, non-inflammatory chronic pain can co-exist in patients with inflammatory RMD, and commonly, fibromyalgia can develop in RA, due to both peripheral and central pain processing.

Therefore, for my study, I used the final analysis-ready dataset, which included 10,584 eligible study participants with complete data on baseline questionnaires and at least one pain data entry. From this dataset, first, I selected study participants with a single chronic pain condition (i.e., no other concomitant chronic pain condition(s)). Specifically, from this cohort with single chronic pain condition, study participants with inflammatory (RA and SpA) and non-inflammatory (OA and chronic widespread pain/fibromyalgia (CWP/FM)) RMDs were selected. The purpose of such cohort selection with single pain condition of different inflammatory and non-inflammatory served for my hypothesis testing, which will allow the exploration of the patterns of daily pain trajectory and variability and to discuss the study findings accordingly with the natural history of pain in each RMD.

Second, from this cohort with inflammatory and non-inflammatory RMDs, only those participants who remained in the study beyond the first 30 days and with pain data for the first 30 days were included as the final dataset for analyses. I censored the study duration to include only the first 30-day period based on results of the previous study of engagement

and retention in *Cloudy with a Chance of Pain*, as outlined in the previous subchapter[253]. This previous study demonstrated the decline in the frequency of data entry after the first one month and therefore, for my study, the choice of 30-day duration for the analyses is a pragmatic solution for having a sufficiently long period to test my study hypotheses and to optimise data completeness[253].

4.6.2 Baseline demographics

During my remote candidature period in the UK, and when I was involved with the research team for the primary analyses of Cloudy with a Chance of Pain study, I was involved in the data cleaning process for the baseline questionnaires, as outlined in the file attached in Appendix A.

For my study, from the data subset of participants with inflammatory (RA and SpA) and non-inflammatory (OA and CWP/FM) RMDs, I extracted their baseline demographics from the final analysis-ready dataset, which included the number of eligible participants for each condition, and their age (in years) and gender. Self-reported pain sites and medications were described as well.

4.6.3 Pain level as the outcome measure of interest

In this exploratory work, I aimed to examine the long-term and short-term day-to-day pain variability in individuals across inflammatory and non-inflammatory RMDs. The daily pain severity was included as the primary outcome of interest for both the multilevel and Markov transition models. Although the pain severity data was reported in an ordinal 5-point scale (1 – no pain, 2 – mild pain, 3 – moderate pain, 4 – severe pain, 5 – very severe pain), I have treated this outcome variable as a continuous variable specifically in the multilevel model, which is an important study limitation to note.

Prior to this, I had explored the nature of the other data items related to the daily pain symptoms reported by study participants in the *Cloudy with a Chance of Pain* study and I described the results as the proportions of each 5-point scale for each data item (fatigue, morning stiffness, impact of pain, sleep quality, time spent outside, waking up feeling tired, physical activity, mood, and wellbeing), as shown in Appendix B.

4.6.4 Statistical methods

For the descriptive data on the baseline demographics, the continuous variables were summarised as mean with standard deviation, and the categorical variables were summarised as absolute numbers and proportions as percentages.

Distribution of pain level

Across all RMDs, for the first 30 days, mean pain scores and proportions for each pain level reporting were visually summarised in a trajectory line plot and a stacked bar chart respectively.

Multilevel model

Specific to my dataset, across different timepoints (level 1), there were repeated measures of pain level of varying intensity for each study participant (level 2) across different RMDs, also known as a hierarchical/multilevel data structure, as illustrated in Figure 4.9.

In detail, these individuals were selected based on their pain conditions and some individuals might share the same type of pain trajectory through time, and more importantly, these individuals with similar characteristics of their pain trajectories might be nested within the same category of either inflammatory or non-inflammatory RMD.

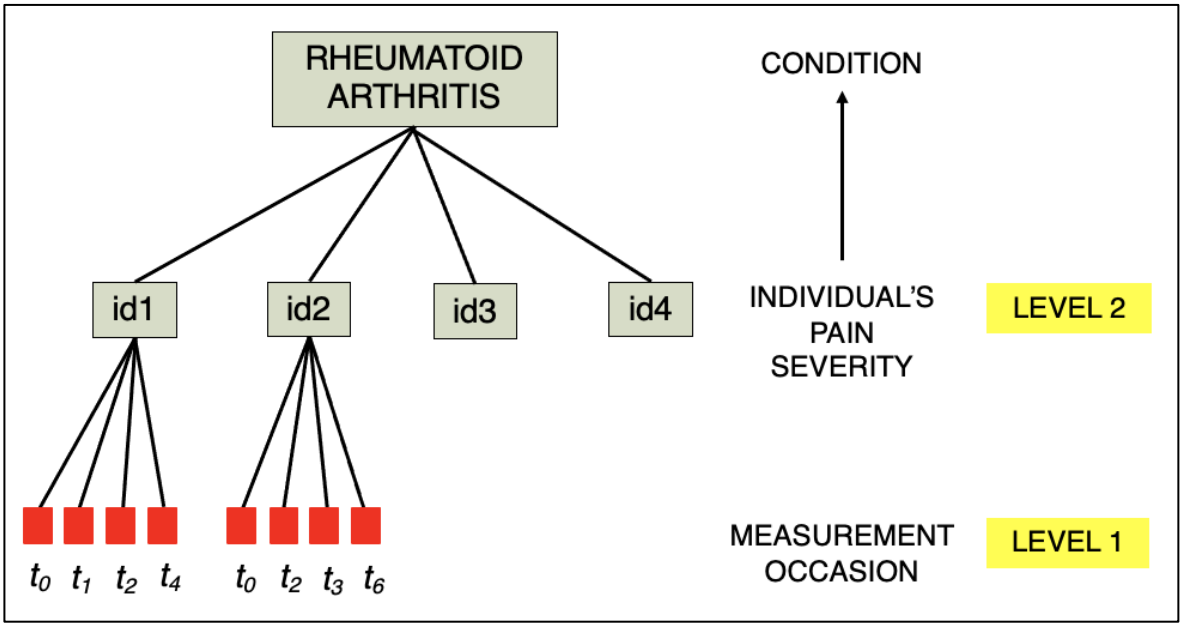


Figure 4.9 An illustrated example of individuals with RA and their corresponding repeated measures of pain (level 2) at different timepoints (level 1) in a multilevel model

Initially, I performed a visual analysis to examine these pain trajectories for individuals with different RMDs, as shown in Figure 4.10. In this illustrated example of a 'spaghetti' plot (essentially a scatterplot with linked data points at each timepoint of measurement) of individuals with RA, I mapped out the population-level mean pain trajectory (highlighted in green line) and the individual-level fitted mean pain trajectories (highlighted in grey lines). As shown in Figure 4.10, as we examine these grey lines in greater detail, there were different types of individual-level pain trajectories, and furthermore, these individuals contributed their pain data at different timepoints. In my exploratory study, with the assumption of normally distributed data, the multilevel model (MLM) (also known as the linear mixed model) is the most suitable statistical model to analyse such longitudinal pain data provided at irregular timepoints of measurement.

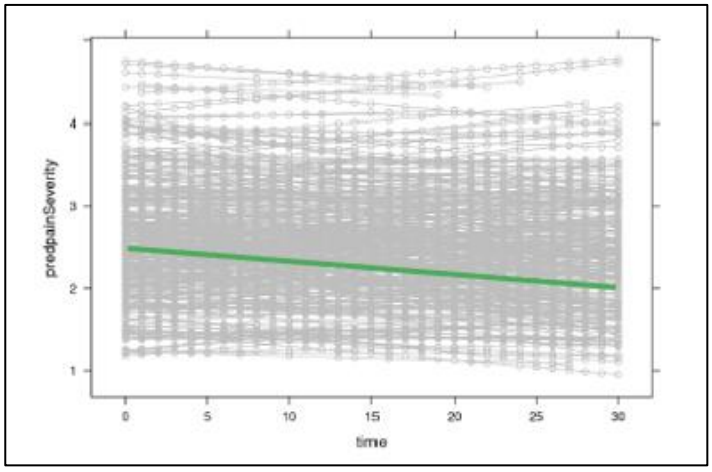


Figure 4.10 An illustrated example of a pain trajectory line plot of individuals with RA over the first 30 days. This plot shows population-level mean pain trajectory pattern over time (in green line) and raw data of individual-specific mean pain trajectories (in grey lines)

In MLM, such hierarchical data structure was applied in the model, with the assumption that an individual's pain responses over time were inter-correlated. This model is used to examine both the between-individual and within-individual data, which contains fixed and random effects[256]. Specifically, the fixed effects refer to the population-level between-individual effects, and the random effects refer to the individual-specific effects. For instance, as seen in Figure 4.10, the population-level intercept is the initial pain level (the start of the green line) and the population-level slope is the average pain trajectory over time. However, to examine the individual-specific pain trajectories in my study (the grey lines in Figure 4.10), as we have multiple individuals with pain data contributed at different timepoints, both random intercept for study participant and a random effect for time (also known as slope) were required in my MLM model.

Additionally, the MLM allowed the residual components at each level of the hierarchical data structure to be examined further. For instance, as shown earlier in Figure 4.9 illustrating a two-level model, this type of MLM allowed the analysis of grouping of pain outcomes across different timepoints within study participants, therefore generating residuals at the pain level (Level 1) and the individual level (Level 2). Within the MLM, the residual variance is stratified into a between-individual component (the variance of the individual-level residuals) and a

within-individual component (the variance of the pain-level residuals) and as a result, the 'unobserved' individual characteristics that affect the pain level over time was defined as the between-individual residuals, also known as the 'individual effects'. Such unobserved variables were instrumental to assess the correlation between outcomes for pain level from the same individual, which was not possible with a traditional fixed effect model. Next, I present a step-by-step MLM approach in my study.

In my exploratory analysis, I initially performed both null (i.e., fixed effect) and random intercept models for comparison of different models examined, and to ensure that the data that I used for my study was applicable using the MLM method. Next, a random slope model was applied to examine the long-term pain variability across individuals with inflammatory and non-inflammatory RMDs, considering the differences in the starting pain level (i.e., varying intercept) and the change in pain over time (i.e., varying slope), as well as simultaneous comparison to the population-averaged pain trajectory within the cohort. This model of choice, although with the assumption of linear change in time, is sufficient to address my study questions, given the pre-defined period of subset of pain severity data selected for my this exploratory work.

The random intercept and slope model formula is outlined below:

$$pain_{ij} = \beta_0 + \beta_1 time_{ij} + u_{0j} + u_{1j} time_{ij} + e_{ij}$$

- $pain_{ij}$ represents the pain response at timepoint measurement *i* (*i* = 1,...,T) for individual *j* (*j* = 1,....,n)
- β_0 represents the population-level intercept (averaged across individuals), which is the expected value of ***pain*** at $time_{ij} = 0$
- β_1 represents the regression slope of *y* on time (which is the growth rate) and in a random slope, it considers a varying growth rate (varying slope) at $time_{ij}$

- u_{0j} represents the individual-specific random intercept for individual *j*, and its variance is the between-individual variance in *pain* after accounting for the linear effect of time
- u_{ij} represents the individual-specific random slope for individual *j* at varying growth rate (varying slope) at $time_{ij}$
- e_{ij} represents the occasion-specific (time-varying) residual, and its variance is the within-individual variance in *pain* after accounting for the linear effect of time

It is important to note that the model assumed that the effect of time on the change in day-to-day pain was linear, which is an important study limitation to note. Additionally, the $time_{ij} = 0$ represented the day of enrolment for the study participants and did not necessarily reflect the timepoint with regards to their RMDs such as disease onset or treatment/intervention.

In this MLM approach, the random effects represent varying distribution of pain trajectories when comparing to the population-level average pain. For instance, some individuals might report higher level of pain and the others might report lower level of pain when compared to the population-level average pain. Additionally, some individuals could have steeper or shallower trajectories than the population average. To include such individual-specific differences in their pain levels (the extent of each individual deviates from the average and the variance between individuals), the model assumed that the individual-specific differences in pain level at baseline $time_{ij} = 0$ followed a normal distribution with a mean of 0 (50% above average and 50% below average). By using this MLM approach, I was able to explore the extent of change in pain over time across different RMDs by examining the correlation between individual-specific intercept and individual-specific slope. For instance, some individuals with higher initial pain level ended with faster rate of change in pain over time (steeper trajectory) and some individuals with lower initial pain level ended

with slower rate of change in pain over time (shallower trajectory) and therefore, resulting in a negative correlation coefficient between the intercept and the slope.

To illustrate my study findings from this MLM approach, the random intercept and slope model output was presented in a slope-intercept plot. As an example for the interpretation of the results, I present a step-by-step description of my study results for each RMD as seen in Figure 4.11, which represents a slope-intercept plot of pain variability in individuals with RA.

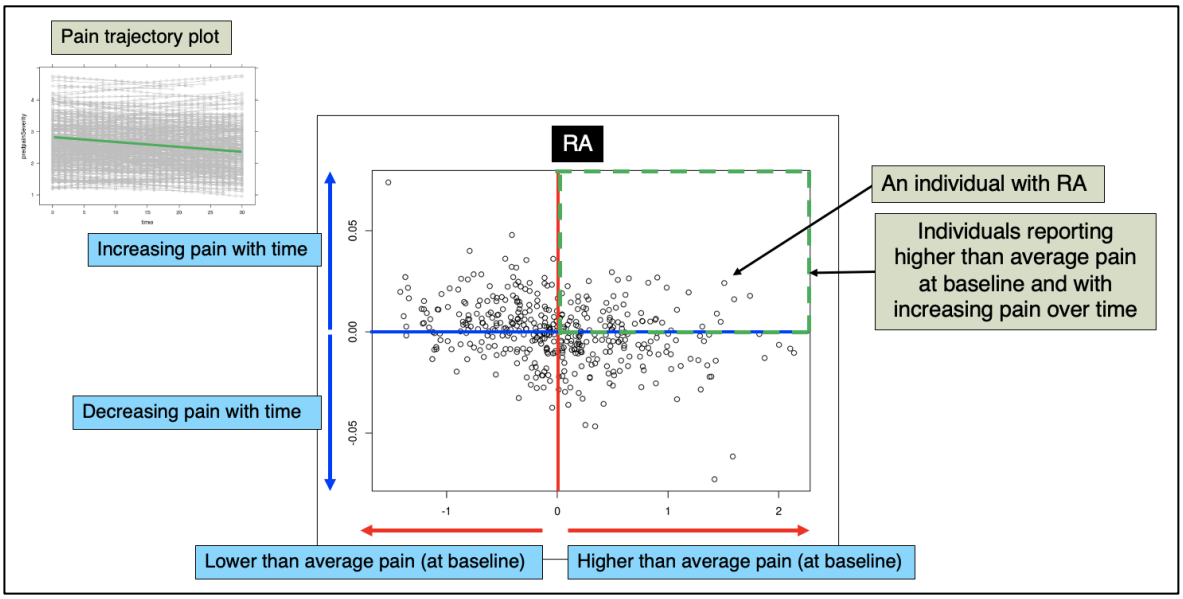


Figure 4.11 An illustrated example of a slope-intercept plot of pain in individuals with RA. In detail, the smaller figure (top left) showed individual pain trajectories (in grey line) a larger figure (in the centre) showed the slope-intercept plot

In Figure 4.11, within the slope-intercept plot, each small circle represented an individual with RA. The red vertical line represented the population-averaged pain score at baseline, with the red arrows pointing to the left or the right as the varying intercepts, representing either lower or higher than average pain at baseline respectively. Horizontally, the blue line represented the population-averaged growth curve of pain with time, with the blue arrows pointing up or down as the varying slopes, representing either increasing or decreasing pain with time respectively. This 4-quadrant slope-intercept plot allowed the examination of individuals with different pain at baseline and with different pain trajectories with time. For

instance, the top right quadrant in the slope-intercept plot (dashed green lined box) included individuals with higher-than-average pain at baseline and with increasing pain over time. In the Results section of this chapter, I will present these slope-intercept plots for each RMD analyses in my study, as well as the predicted population-level pain trajectories for each RMD.

This MLM modelling analysis was performed using the *lme4* package in R Statistical Software (v3.6.2; R Core Team 2021)[257]. The statistical code scripts and the MLM model output (null, random intercept, random slope) and model selection were described in detail in Appendix C.

Markov transition model

Given the 5-point ordinal score of the pain level in the study, assuming that the future pain level is dependent on the current pain level, I applied a single sequence, first-order Hidden Markov model (HMM) or Markov transition model in examining the short-term pain variability in individuals with inflammatory and non-inflammatory RMDs.

In general, HMM is based on the Markov chain process, first introduced by Andrei Andreyevich Markov (1856-1922), which describes a stochastic transition process of which the future state is dependent on the present state, considering the periodicity and recurrence of the states[258, 259]. Such probabilistic analytical approach is appropriate to examine the likelihood of change in pain states in my study, and as I aimed to examine the effect of lag difference in time of 1 (i.e., comparison of current pain state to the pain state one day prior), I applied the first-order HMM in my analysis.

The model output was presented in a heatmap transition plot. As an example for the interpretation of the results using the HMM approach, I present a step-by-step description of my study results for each RMD as seen in Figure 4.12, which represents a heatmap transition plot of one single level of day-to-day transition between pain states (short-term

pain variability) in individuals with CWP/FM. In this heatmap plot, only the 'very severe' pain level transition probability was described.

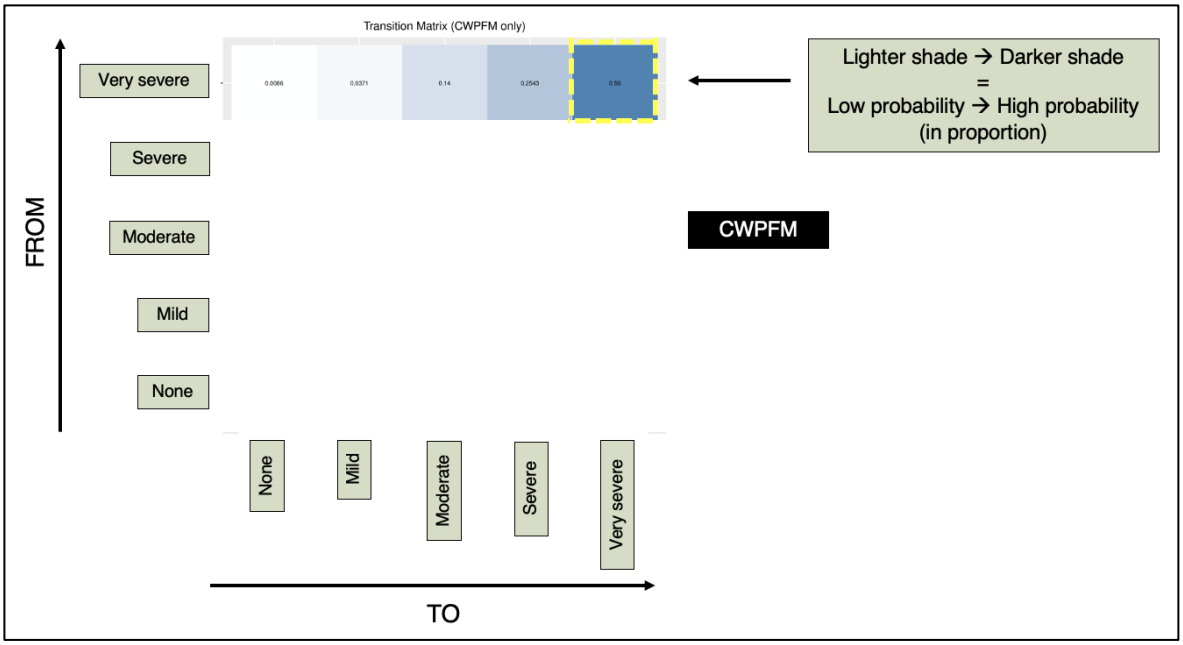


Figure 4.12 An illustrated example of a heatmap transition plot of a single transition matrix between pain states in individuals with CWP/FM. In this heatmap plot with varying colour shades, lighter shade indicates low probability of staying in the same pain state and darker shade indicates high probability of staying in the same pain state.

In Figure 4.12, focusing on a single transition matrix of 1-5 (none-very severe) ordinal pain level, the lighter shade corresponded to a lower probability of transitioning between pain states and the darker shade corresponded to a higher probability of transitioning between pain states. In this example of individuals with CWP/FM, there were higher proportions of individuals in the 'very severe' pain state, having transitioned from a 'very severe' pain state prior. In the Results section of this chapter, I will present these heatmap transition plots for each RMD analyses in my study.

This analysis was performed using the *seqHMM* package in R Statistical Software (v3.6.2; R Core Team 2021)[260]. The statistical code scripts and the Markov transition model output were described in detail in Appendix C.

Pain fluctuation

In addition to the HMM analysis, the day-to-day pain fluctuations (the magnitude of either increase or decrease in pain level) were also summarised across all conditions as proportions, which I presented the results in stacked bar chart. Based on a previous pain flare definition analysis on Cloudy dataset by Beukenhorst and colleagues, I presented my pain fluctuation analysis as the absolute change and significant change of pain from yesterday – either an increase in pain by +2 (e.g., none → moderate) and +4 (e.g., none → very severe) or a decrease in pain by -2 (e.g., moderate → none) and -4 (e.g., very severe → none)[261]. No change in pain level from yesterday was specified by a value of 0 in my analysis.

4.7 Results

Baseline demographics for individuals across different RMDs were described in Table 4.2. Of the total of 1,189 included study participants, they were predominantly female and aged between 40 and 60 years old. A younger age groups of participants were seen in those with SpA and CWP/FM.

Higher proportions of reported pain sites resembling the typical patterns of joint involvement for RA, SpA and OA were observed (highlighted in bold for each condition), with paracetamol and NSAIDs being commonly used across these conditions (highlighted in bold for each condition) (Table 4.2). In contrast, higher proportions of varying reported pain sites were seen in those with CWP/FM, and similarly, with a wider range of pain medications being used more commonly (highlighted in bold) (Table 4.2).

	Rheumatoid arthritis (RA)	Spondyloarthritis (SpA)	Osteoarthritis (OA)	Chronic widespread pain/fibromyalgia (CWPFM)
N of participants	425	100	409	255
Female gender (%)	354 (83.3)	65 (65.0)	340 (83.1)	228 (89.4)
Age: mean (SD)	49.17 (11.70)	44.90 (11.96)	57.01 (10.65)	41.44 (10.89)
PAIN SITE				
Head (%)	21 (4.9)	9 (9.0)	13 (3.2)	70 (27.5)
Face (%)	6 (1.4)	2 (2.0)	4 (1.0)	37 (14.5)
Mouth/jaw (%)	59 (13.9)	13 (13.0)	14 (3.4)	71 (27.8)
Neck/shoulder (%)	**231 (54.4)**	**64 (64.0)**	151 (36.9)	**157 (61.6)**
Back (%)	122 (28.7)	**88 (88.0)**	186 (45.5)	**165 (64.7)**
Abdomen (%)	20 (4.7)	11 (11.0)	15 (3.7)	64 (25.1)
Hip (%)	175 (41.2)	**57 (57.0)**	207 (50.6)	**150 (58.8)**
Knee (%)	**251 (59.1)**	45 (45.0)	**282 (68.9)**	**149 (58.4)**
Hands (%)	**324 (76.2)**	40 (40.0)	**215 (52.6)**	**132 (51.8)**
Feet (%)	**267 (62.8)**	35 (35.0)	147 (35.9)	**112 (43.9)**
Multi-site (%)	197 (46.4)	41 (41.0)	67 (16.4)	**169 (66.3)**
All over body (%)	45 (10.6)	5 (5.0)	13 (3.2)	**121 (47.5)**
MEDICATION				
No analgesia (%)	38 (8.9)	7 (7.0)	35 (8.6)	14 (5.5)
Paracetamol (%)	**182 (42.8)**	**38 (38.0)**	**220 (53.8)**	**118 (46.3)**
NSAIDs (%)	**270 (63.5)**	**75 (75.0)**	**261 (63.8)**	**133 (52.2)**
Other simple analgesia (%)	124 (29.2)	22 (22.0)	114 (27.9)	86 (33.7)
Weak opioids (%)	93 (21.9)	25 (25.0)	101 (24.7)	**107 (42.0)**
Strong opioids (%)	30 (7.1)	13 (13.0)	26 (6.4)	29 (11.4)
Neuropathic pain agent (%)	22 (5.2)	13 (13.0)	31 (7.6)	**108 (42.4)**
Other analgesia (%)	10 (2.4)	1 (1.0)	25 (6.1)	47 (18.4)

Table 4.2 Baseline demographics across individuals with inflammatory (RA and SpA) and non-inflammatory (OA and CWP/FM) RMD

From a total of 23,470 daily pain scores in this cohort, and with a median days of data entry of 26 (IQR 21-30), participants with CWP/FM had the highest mean pain score of 3.04±1.03 and the highest proportion of moderate to very severe pain reporting, with nearly 70% of these participants reporting moderate to very severe pain (Figures 4.13 and 4.14).

Compared to those with RA, who had the least mean pain score of 2.50±0.98 and the lowest proportion of moderate to very severe pain reporting (46%), the mean pain scores were slightly higher in those with SpA and OA (2.73±0.98 and 2.60±0.95 respectively), corresponding to their proportions of moderate to very severe pain reporting (57% and 51% respectively) (Figures 4.13 and 4.14).

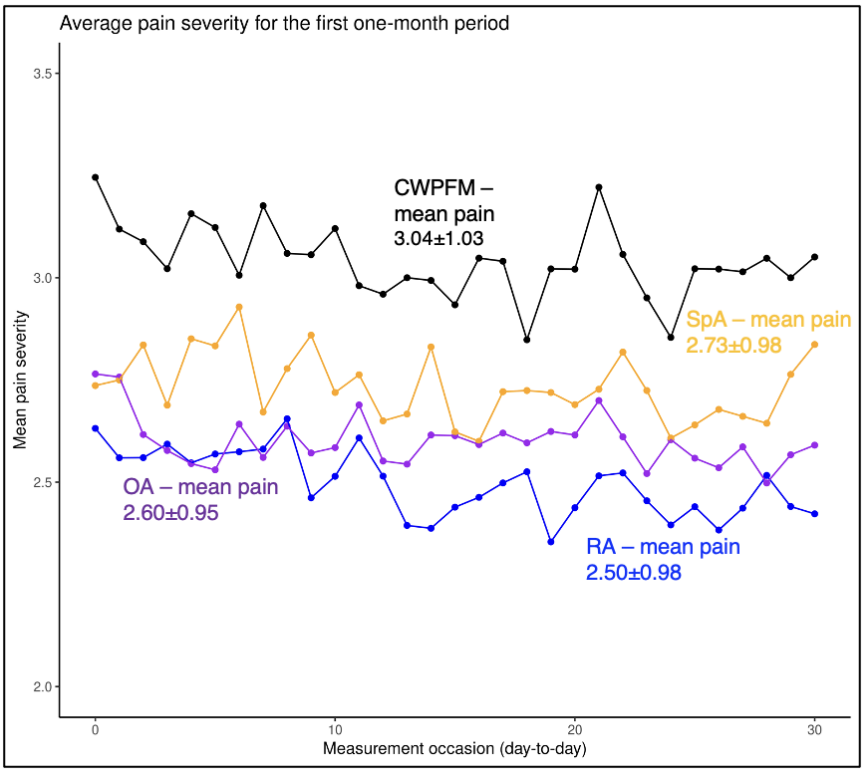


Figure 4.13 Mean pain scores for the first 30-day period for individuals with RA, SpA, OA, and CWP/FM

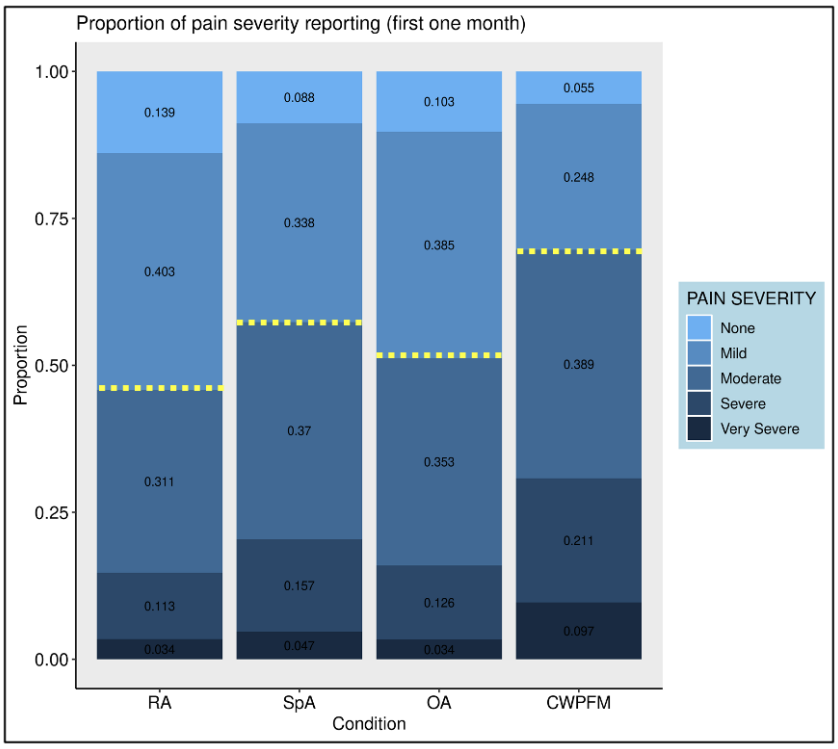


Figure 4.14 Proportion of pain severity reporting for the first 30-day period for individuals with RA, SpA, OA, and CWP/FM

In the MLM, from the population level, as shown in Figure 4.15, across all conditions, there were steeper time-based improvements in pain for individuals reporting higher initial pain scores.

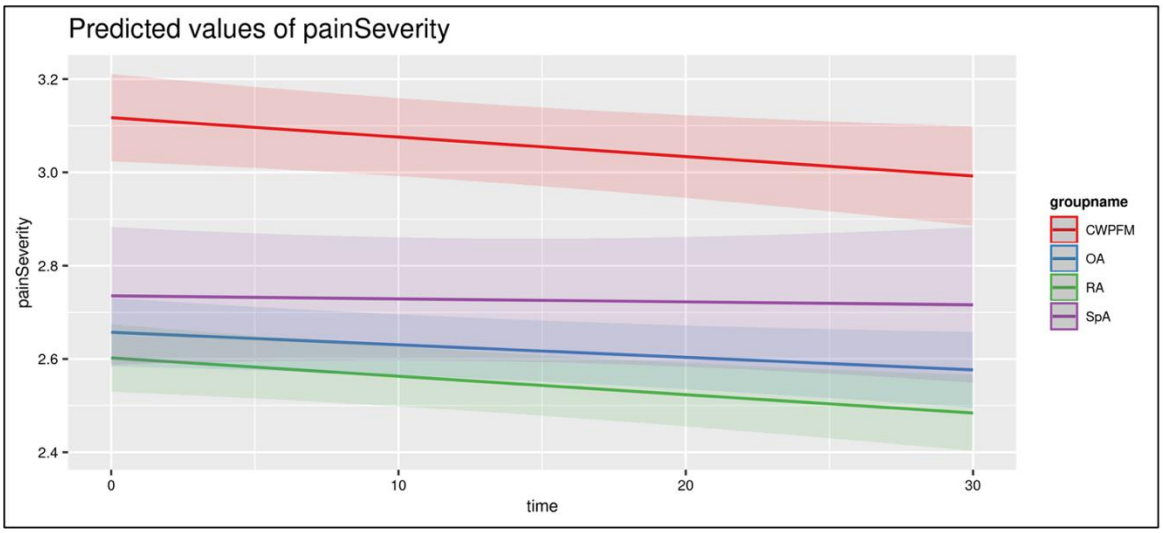


Figure 4.15 Predicted population-level mean pain trajectories for individuals with RA, SpA, OA, and CWP/FM

From the individual level, as illustrated in the slope-intercept plots, across all conditions, the long-term change in pain was significantly different between individuals (represented by the circles in the slope-intercept plot for each condition) (Figure 4.16). Negative correlation coefficients were seen across all conditions from the random slope model, which corresponded to the time-based improvements in pain for individuals reporting higher initial pain scores (Figure 4.16).

In the Markov transition model, as shown in Figure 4.17, across all conditions, similar 'diagonal' patterns of darker shades (represented by the diagonal arrows in the heatmap plot) were observed, which indicated the highest probability of staying in the same pain states and minimal variation in day-to-day change in pain states. Of note, in CWP/FM, 56% of those in the 'very severe' pain state stayed in that pain state with minimal day-to-day variation (indicated by the highlighted area in the heatmap plot for CWP/FM on the bottom right of Figure 4.17).

Across all conditions, there were no day-to-day fluctuations in pain state in 50% of days, although an increase in pain level was observed in 25% of days, with a 2-point increase in pain level or more observed in 4% of days (Figure 4.18).

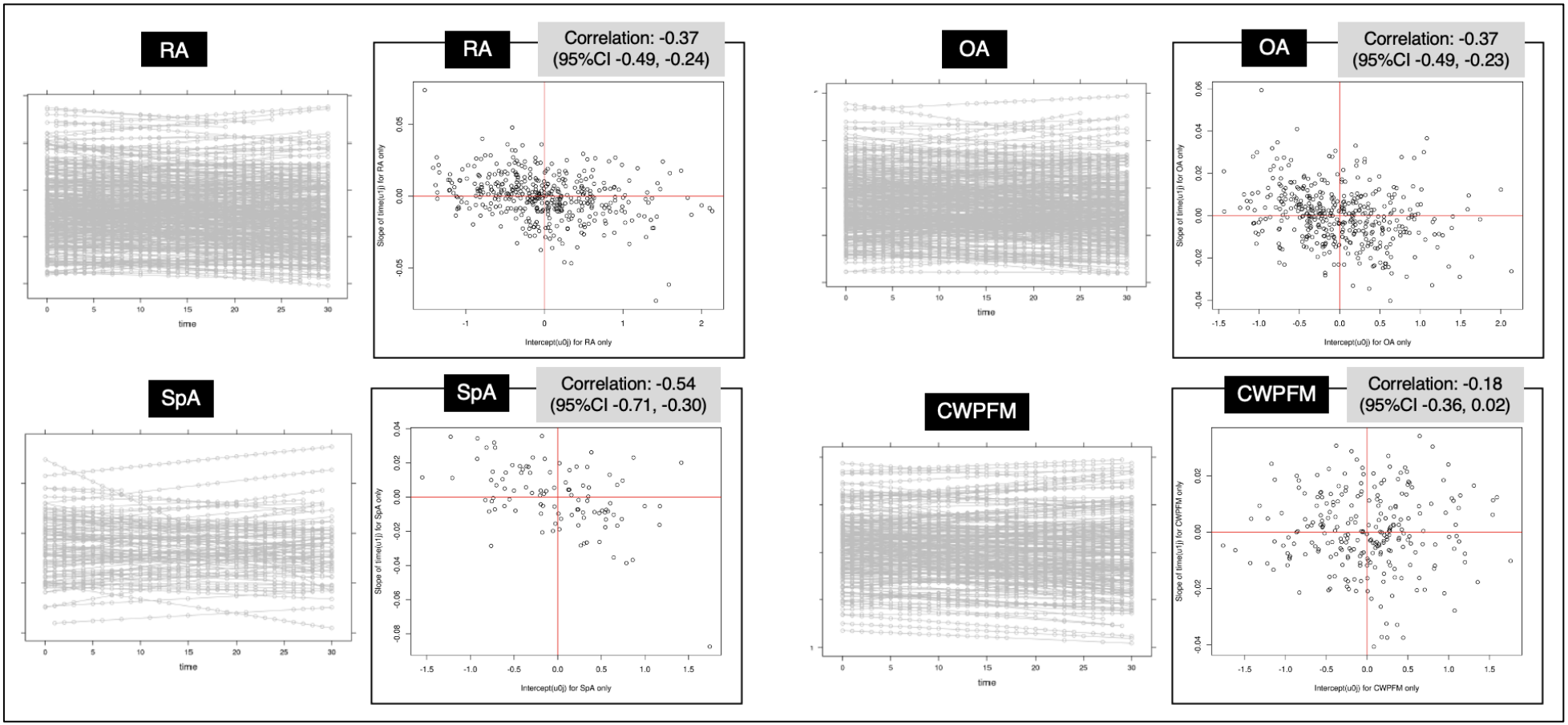


Figure 4.16 Pain trajectory plot (left panel for each condition) and slope-intercept plot (right panel for each condition) with correlation coefficients of long-term change in pain over time for individuals with RA, SpA, OA, and CWP/FM. Of note, negative correlation coefficients were seen across all RMDs.

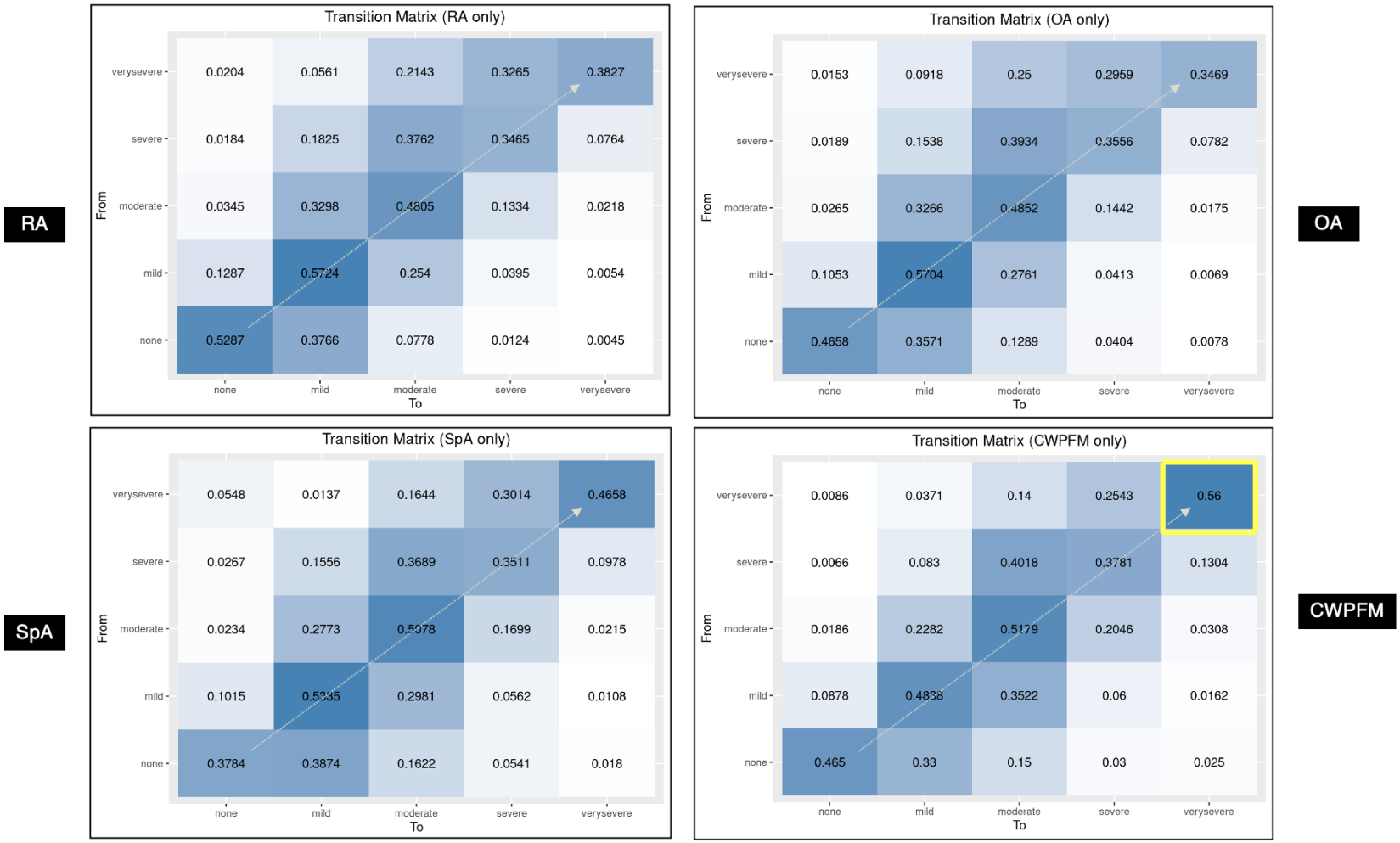


Figure 4.17 Heatmap plot with transition probabilities indicated in each matrix for individuals with RA, SpA, OA, and CWP/FM. Similar diagonal patterns of darker shades were seen across all conditions (arrow) and in CWP/FM, 56% stayed in the 'very severe' pain state from yesterday's pain state (yellow box)

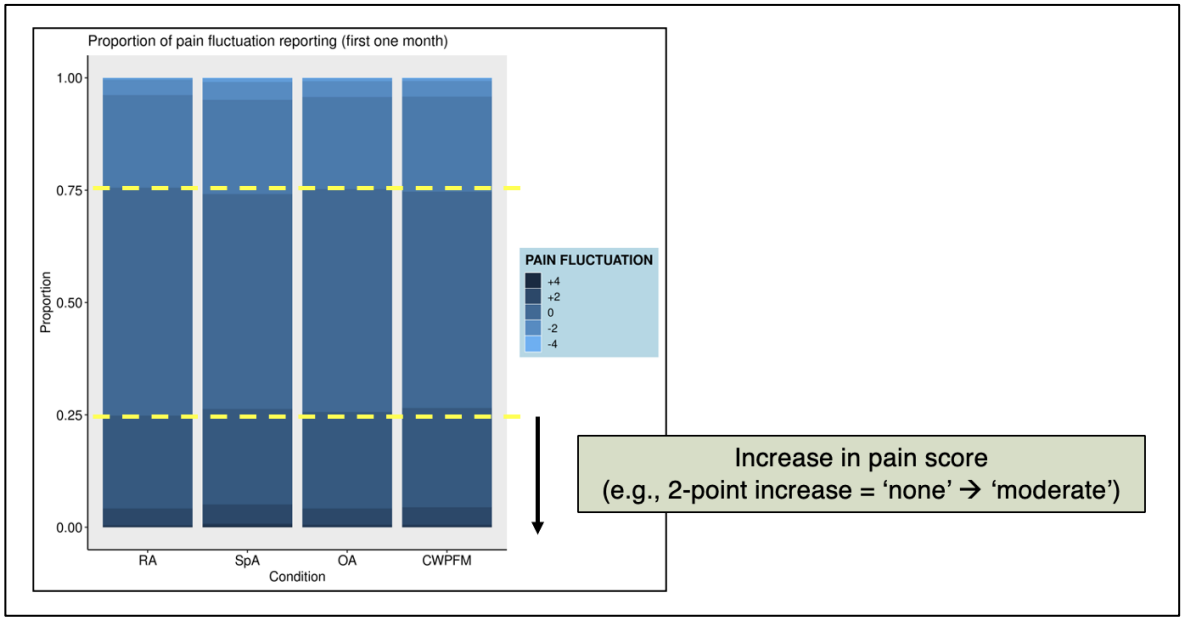


Figure 4.18 Proportion of pain fluctuation reporting for the first 30-day period for individuals with RA, SpA, OA, and CWP/FM

4.8 Discussion

In *Cloudy with a Chance of pain* study, by harnessing citizen scientist-generated daily pain symptom data using smartphones, I was able to examine patterns of day-to-day pain variability in individuals with inflammatory (RA and SpA) and non-inflammatory (OA and CWP/FM) RMDs. These analyses of pain variability were achieved using intensive longitudinal methods, of which I applied both multilevel model and Markov transition model to examine long-term and short-term pain variability respectively. Three important key study findings were identified in these analyses, which are next discussed in detail.

4.8.1 Key finding 1: Highest mean pain level and highest proportions of moderate-very severe pain reporting were seen in individuals with CWP/FM

Using more than 23,000 daily pain scores over the first 30-day period, and excluding other confounding comorbid pain condition, individuals with CWP/FM reported both the highest

mean pain level of 3.04±1.03 and proportions of moderate-very severe pain reporting (70%), followed by those with SpA, OA, and RA (mean pain level of 2.73±0.98, 2.60±0.95 and 2.50±0.98 respectively; proportions of moderate moderate-very severe pain reporting of 57%, 51% and 46% respectively).

Evidently in CWP/FM, as seen in the Markov transition model, more than half of those in the 'very severe' pain state remained in that pain state with significantly minimal day-to-day variation in pain states. Such findings represent the common symptomatology experienced by individuals with either chronic widespread pain or fibromyalgia[48, 262]. Amplification of pain in FM, typically coming from both nociceptive pain processing and central hypersensitisation, is a unique feature that distinguish these individuals from the others in this study[48, 262].

4.8.2 Key finding 2: Patterns of pain improvement in individuals with higher initial pain level were seen across all conditions

In this study, based on the pain trajectory plots generated across different single-disease RMDs, despite starting at higher pain level, patterns of improvement in pain were seen across all conditions. In research involving repeated data analysis, regression to the mean is a common statistical phenomenon of observing true change in the natural progression of repeated measurements, especially in dataset with random measurement error and also with random fluctuations of the follow-up outcome measure derived from a subset of data chosen using a baseline value[263, 264]. In my study, in the context of using a mobile app to enter data items relating to pain, it is plausible that the included study participants were likely to have worst pain at the time of joining the study, and their pain might start to settle as time progressed. In this case, regression to the mean is a plausible explanation of these downtrending pain trajectories, given that participants could join in the Cloudy study at any timepoint regardless of the timing of their disease process.

4.8.3 Key finding 3: The volatility of changing pain states was comparable across all conditions

By using all the daily data points in the Markov transition analysis, we observed that the probability was highest in those remaining in the same pain states from yesterday, indicating minimal day-to-day variation in pain states. Similarly, from the day-to-day pain fluctuation plots, we did not see frequent absolute change in pain level from yesterday, either 2-point or 4-point increase and decrease in pain. The volatility of change in pain states was relatively stable and comparable across all conditions, suggesting the possibility of no difference in flare periods. Of note, in CWP/FM, more than 50% remained in the same 'very severe' pain state with minimal day-to-day transition, indicating the magnitude of chronicity of high pain in this condition.

4.8.4 Strengths and limitations

In this study, despite restricting the dataset to a 30-day period for these analyses, patterns of day-to-day pain variability were well described using both modelling approaches – the MLM and the Markov transition model. This exploratory study of pain variability forms the basis of concise understanding in the dynamic process of pain response across different RMDs using intensive longitudinal methods. In addition, the MLM approach allows the integration of both between-individual and within-individual components in the analysis of pain trajectories.

Nevertheless, there were some limitations in this study. First, female predominance was seen in my study cohort, and therefore, the study results were more likely to represent the female cohort with the RMDs analysed. Second, *Cloudy with a Chance of Pain* study was designed as a citizen science research study and objective measures of disease activity were not available. As a result, the severity of the disease in this study is largely unknown and it was not possible to ascertain the corresponding disease activity to the individual's pain trajectory or changing pain states. Nevertheless, this smartphone study has shown us

the success of a large-scale recruitment of study participants and prospective daily patient-generated health data collection using the 'Cloudy' mHealth app for symptom capture. Such real-time symptom tracking in mHealth is a promising future for self-management and self-monitoring of pain in RMDs, with many possibilities of early identification of concerning flare periods or worsening pain in real-time and a window of opportunity to intervene, when correlating these symptom data with clinical assessment and other objective disease activity measures[190]. Third, the study cohort selection for my analysis was considered near-complete cases, with an average of 2 out of 30 days of missing data per individual. Therefore, in my study analysis, study findings were limited to a short period, with a balance to minimise data missingness. Future study may further validate these study findings by using the full dataset along with multiple imputations in handling the missing data. Fourth, as highlighted in the Methods section in this Chapter, in the MLM, the effect of time on day-to-day change in pain was considered as linear or fixed. Such limitation of assuming a stable or linear rate of change is worth highlighting, as this may not be necessarily true for different rates of change in pain that may vary at different timepoints even for the same individual. Alternative repeated measures analytical approaches such as polynomial trends in time, regression spline, orthogonal polynomial, and fractional transformations have been considered in delineating the non-linear effect of time on changes in pain, which is beyond the scope of this exploratory study[233, 265-270]. They are, however, being explored further by others using this dataset. Fifth, due to the simplistic Markov transition modelling approach in my study, other contextual factors influencing the change in pain states were not examined (e.g., sleep, mood, physical activity, medication use). Local within-individual probability of pain state transition was not examined in this study, and alternative Markovian-based model of Bayesian dynamic time series may be required.

4.8.5 Contribution as co-author in the primary publications

In the Appendix D of this thesis, I have included three primary publications derived from Cloudy with a Chance of Pain. I was included as one of the co-authors for these publications

as I was involved in the work leading up to these publications during my initial remote candidature period. Apart from contributing to the study analysis, interpretation of the study results and the drafting of these publications, I also assisted in preparing the supplementary materials with Dr Anna Beukenhorst, Dr Belay Birlie Yimer and Dr Jamie Sergeant.

These publications include:

- Dixon, W.G., Beukenhorst, A.L., Yimer, B.B., Cook, L., Gasparrini, A., El-Hay, T., Hellman, B., James, B., Vicedo-Cabrera, A.M., Maclure, M., Silva, R., Ainsworth, J., **Pisaniello, H.L.**, House, T., Lunt, M., Gamble, C., Sanders, C., Schultz, D.M., Sergeant, J.C., McBeth, J. How the weather affects the pain of citizen scientists using a smartphone app. NPJ Digital Medicine, 2019; 2:105. doi.org/10.1038/s41746-019-0180-3 [**Published**]
- Schultz, D.M., Beukenhorst, A.L., Yimer, B.B., Cook, L., **Pisaniello, H.L.**, House, T., Gamble, C., Sergeant, J.C., McBeth, J., Dixon, W.G. Weather Patterns Associated with Pain in Chronic-Pain Sufferers. Bulletin of the American Meteorological Society (BAMS), 2020; 101(5). doi/org/10.1175/BAMS-D-19-0265.1 [**Published**]
- Yimer, B.B., Schultz, D.M., Beukenhorst, D.M., Lunt, M., **Pisaniello H.L.**, House, T., Sergeant, J.C., McBeth, J., Dixon, W.G. Heterogeneity in the association between weather and pain severity among patients with chronic pain: a Bayesian multilevel regression analysis. Pain Reports, 2022;7doi.org/10.1097/PR9.0000000000000963 [**Published**]

I have highlighted this valuable learning experience in epidemiology and research skills in a lay summary presented on the Arthritis Australia website (previous award recipient) following the receipt of the overseas fellowship award (Ken Muirden Overseas Training Fellowship funded by the Australian Rheumatology Association, ARA) which funded this opportunity (Appendix E).

4.9 Summary

Cloudy with a Chance of Pain was a successful citizen-scientist study examining the weather-pain relationship using smartphones. In this study analysis I presented using a subset of the *Cloudy with a Chance of Pain* data over the first 30-day period, there were several key study results generated when examining the day-to-day pain trajectory and pain variability among individuals across different RMDs.

Main findings include:

- In comparisons to individuals with SpA, OA and RA, highest mean pain level and highest proportions of moderate-very severe pain reporting were seen in individuals with CWP/FM
- In those with CWP/FM, those with a 'very severe' pain state remained in the same pain state from yesterday and had minimal change in day-to-day pain state
- In the multilevel model, patterns of pain improvements in individuals, despite starting at higher pain level, were seen across all conditions (RA, SpA, OA, and CWP/FM), with regression to the mean as a plausible explanation of this study finding
- In the Markov transition analysis, highest probability of staying in the same pain states from yesterday was seen across all conditions (RA, SpA, OA, and CWP/FM). Similarly, from the day-to-day pain fluctuation plots, there was minimal absolute change in pain level from yesterday, defined as 2-point or 4-point increase and decrease in pain. These findings suggest that the volatility of changing pain states was minimal and comparable across all conditions (RA, SpA, OA, and CWP/FM), suggesting no difference in flares
- This study provides the opportunity to use intensive longitudinal methods to analyse day-to-day pain variability in RMDs by harnessing large dataset of pain symptoms, albeit the selection of the first 30-day study period for data completeness and evidence of participant attrition in the previous study

Summary of Section 2

Section 2 presents the results of chapters 3 and 4, which address the second research question in this thesis in terms of discussing the benefits and challenges of harnessing digitalised health-related data and the use of intensive longitudinal methods in examining long-term and short-term day-to-day pain variability across different RMDs using a daily pain data derived from a UK-based smartphone study, *Cloudy with a Chance of Pain*.

Specifically, the narrative review in chapter 3 highlights the importance of governance in data protection, consent, and ownership as we continue to see the ever-rising digitalised health-related data collection, use, and application in clinical practice and research.

In chapter 4, using both multilevel model and Markov transition model to analyse the day-to-day pain variability across different RMDs for a subset data of *Cloudy with a Chance of Pain*, study results have shown patterns of pain improvement for individuals with higher initial pain scores and minimal change in day-to-day pain states across all conditions, suggesting minimal day-to-day pain fluctuation. Such study results highlight the advantages of using this modelling approach to examine the extent of pain variability over time, however, the magnitude of response shift of pain over time is not clear. The original intention of Cloudy with a Chance of Pain study is to examine the association of pain and weather, and therefore, the dataset has limitations with further exploration of important pain-related health outcome measures. Additionally, persistent pain in those with inflammatory RMDs may not be evident in the current analysis presented in Chapter 4, given the lack of supporting disease-related outcome measures. Nevertheless, the analytical approach and key study findings from this work presented in Chapter 4 have formed the stepping stone to my pain trajectory work in RA, which are presented in Section 3 of this thesis. In the next section, the impact of pain trajectories in RA are closely examined, in the setting of identifiable subgroups of different pain trajectories using clustering methods. In Chapter 5, the trajectory of DAS28 in an early RA cohort is examined using the clustering method,

followed by analyses of the identified trajectory subgroups of the treatment characteristics, disease progression, and the sociodemographic factors. In Chapters 6 and 7, in an established RA cohort derived from a national biologic DMARD registry, I extend the analysis into a multi-trajectory analysis of different pain-related health outcomes and the associations with treatment characteristics, sociodemographic factors, comorbidities, hospitalisations and mortality rates.

Section 3: The Attributable Burden of Persistent Pain in Rheumatoid Arthritis (RA)

Overview of Section 3

Section 3 of this thesis consist of three chapters, chapters 5, 6, and 7, which address the final research question of this thesis in identifying at-risk individuals with persistent pain in RA, and in exploring the impact of persistent pain on medication use and important health outcomes in RA, such as hospitalisation events, mortality risk, and causes of death.

In chapter 5, the study aimed to utilise the DAS28-P index, a derived proportion of the patient reported components within the DAS28, to identify distinct subgroups of individuals with early RA having different trajectories of the DAS-28P index over 1 year. By using these trajectory subgroups, predictors of treatment response to DMARD therapy were explored in this cohort.

Chapters 6 and 7 present the work done using the Australian Rheumatology Association Database (ARAD), a national registry of study participants with inflammatory arthritis and monitoring of treatment safety and patient reported outcomes. In chapter 6, first, using a discrete mixture modelling approach, distinct multi-trajectory subgroups of five different self-reported pain-related health outcomes were identified over a 15-year follow-up period. Next, using these trajectory subgroups as predictors, the study aimed to examine the associations of these trajectory subgroups with their corresponding baseline demographics, sociodemographic indicators, comorbidities, and medication use. In chapter 7, the study presented is an extension of chapter 6. Using similar clustering method of identifying distinct multi-trajectory subgroups of pain-related health outcomes, the study presented in this chapter aimed at examining the associations of these identified trajectories of pain-related health outcomes with hospitalisation events, mortality risk, and causes of death.

Chapter 5: DAS28-P Index as a Discriminatory Measure of Treatment Response in Early RA

5.1 Preface

The research work presented in this chapter addresses the research question of utilising disease activity score, and in this case, of the DAS28-P index, to identify distinct trajectories over 1 year in an early RA cohort. By using these trajectory subgroups as predictors, the effect of DAS28-P index on treatment response to DMARD therapy was examined.

This study is presented in a manuscript format in this chapter, as it has been published in a peer-reviewed journal, BMC Rheumatology, as:

Pisaniello, H.L., Whittle, S.L., Lester, S., Menz, F., Metcalf, R., McWilliams, L., Hill, C.L., Proudman, S. Using the derived 28-joint disease activity score patient-reported components (DAS28-P) index as a discriminatory measure of response to disease-modifying anti-rheumatic drug therapy in early rheumatoid arthritis. BMC Rheumatology, 2022; 6:67. doi.org/10.1186/s41927-022-00299-3

The Statement of Authorship is included. To end, I present the chapter summary from this study.

5.2 Dataset: Early Rheumatoid Arthritis Cohort

The Early Rheumatoid Arthritis Cohort is a prospective cohort derived from a double-blind, randomised controlled trial (RCT) in examining the effects of fish oil in RA, and the original study protocol is outlined here in detail[271].

5.2.1 Study participants

Consecutive patients presenting with recent-onset polyarthritis to the Early Arthritis Clinic at the Royal Adelaide Hospital (RAH), South Australia, Australia were screened. The eligibility for study participation was performed by the principal investigator between September 2001 and December 2008.

Inclusion criteria:

- Aged 18 years or older
- Diagnosis of RA according to the 1987 revised American College of Rheumatology (ACR) criteria[272]
- Symptoms of polyarthritis of <12 months' duration, with a swollen joint count (SJC) ≥3, a tender joint count (TJC) ≥6, an erythrocyte sedimentation rate (ESR) >28mm/hour, and/or C-reactive protein (CRP) >10mg/dL
- DMARD-naive
- Written informed consent was obtained

Exclusion criteria:

- DMARD use other than anti-malarials
- Use of anti-malarials for more than 1 month
- Recent seroconversion to parvovirus, Ross River, Barmah Forest, or rubella viruses
- History of positive anti-nuclear antibody (ANA) with titre ≥1:320
- History of positive hepatitis B, hepatitis C, or human immunodeficiency virus (HIV)

- Known sensitivity to methotrexate, sulphasalazine, or hydroxychloroquine
- History of systemic disease likely to increase the risk of toxicity to 1 or more of these DMARDs

Study approval was obtained from the RAH Research Ethics Committee. Eligible study participants provided their written informed consent. This study was registered under the Australian New Zealand Clinical Trials Registration number ACTRN12613000579796.

5.2.2 Study screening process

A total of 187 consecutive patients attending the Early Arthritis Clinic were screened for eligibility between September 2001 and December 2008. Of 187 patients screened, 12 eligible patients originated from The Queen Elizabeth Hospital clinics and the remaining 175 eligible patients were recruited from the RAH clinics. All 187 patients were evaluated by the RAH metrologist who did not change for the whole study duration.

Of these 187 participants, 32 participants did not fulfil the eligibility criteria due to various reasons: failure to fulfil the ACR criteria for RA, were not DMARD-naïve, absence of active disease, and inability to provide informed consent. An additional 15 participants fulfilled the eligibility criteria but declined to participate in the study due to various reasons: through choice, planned pregnancy in the next 2 months, or died before study enrolment. In total, 140 participants were subsequently randomised to receive either high dose fish oil or low dose fish oil. Figure 5.1 illustrates the screening process derived from the original paper[271].

Study cohort selection specific to my study presented in this chapter is discussed in detail in the published article.

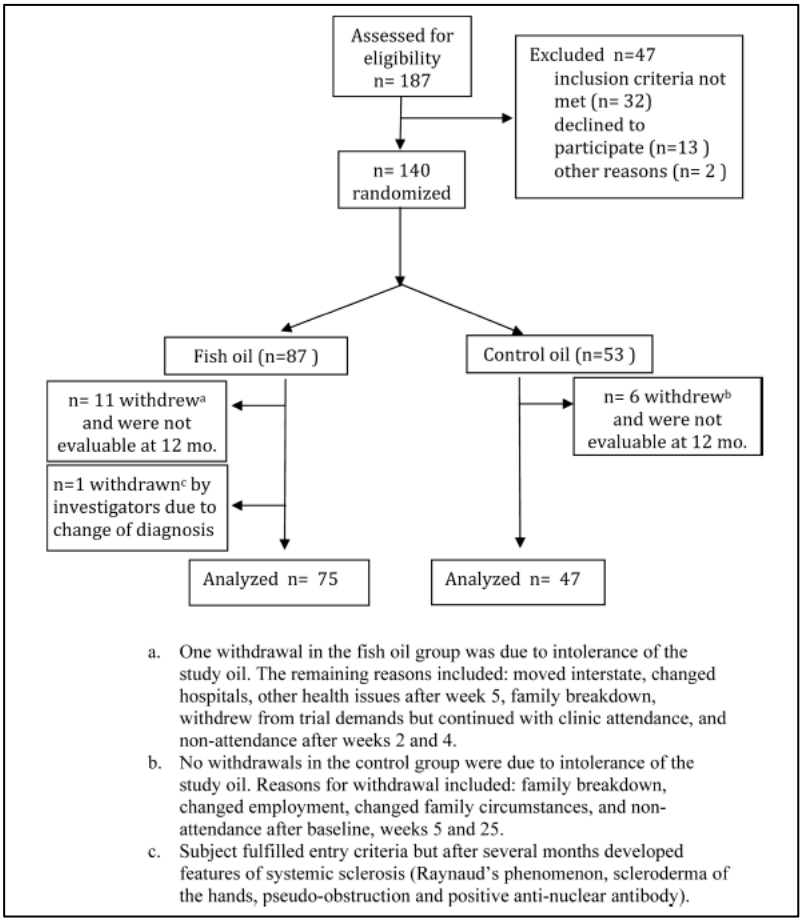


Figure 5.1 Recruitment and eligible study cohort selection process for Early Arthritis Clinic in a flow chart

Source: Adapted from Proudman SM, James MJ, Spargo LD, et al[271] (used with permission under Creative Commons Attribution License)

5.3 Statement of Authorship

Statement of Authorship

Title of paper	Using the derived 28-joint disease activity score patient-reported components (DAS28-P) index as a discriminatory measure of response to disease-modifying anti-rheumatic drug therapy in early rheumatoid arthritis
Publication status	Published
Publication Details	Pisaniello, H.L., Whittle, S.L., Lester, S., Menz, F., Metcalf, R., McWilliams, L., Hill, C.L., Proudman, S. Using the derived 28-joint disease activity score patient-reported components (DAS28-P) index as a discriminatory measure of response to disease-modifying anti-rheumatic drug therapy in early rheumatoid arthritis. BMC Rheumatology, 2022; 6:67 doi.org/10.1186/s41927-022-00299-3 Impact Factor: 2.451

Principal Author

Name of principal author	Huai Leng Pisaniello
Contribution to the Paper	Contributed to the study design and conception, and data acquisition and management. Contributed to the data analysis and interpretation of study results. Contributed to the preparation, drafting and revision of the manuscript for publication.
Overall percentage (%)	75%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 8 April 2023

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. The candidate's stated contribution to the publication is accurate (as detailed above);
ii. Permission is granted for the candidate in include the publication in the thesis; and
iii. The sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Samuel L. Whittle
Contribution to the Paper	Contributed to the study design and conception, and data acquisition. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 18/04/2023

Name of Co-Author	Susan Lester
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 18/04/2023

Name of Co-Author	Fiona Menz
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the data analysis. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 15, 2023

Name of Co-Author	Robert Metcalf
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 20, 2023

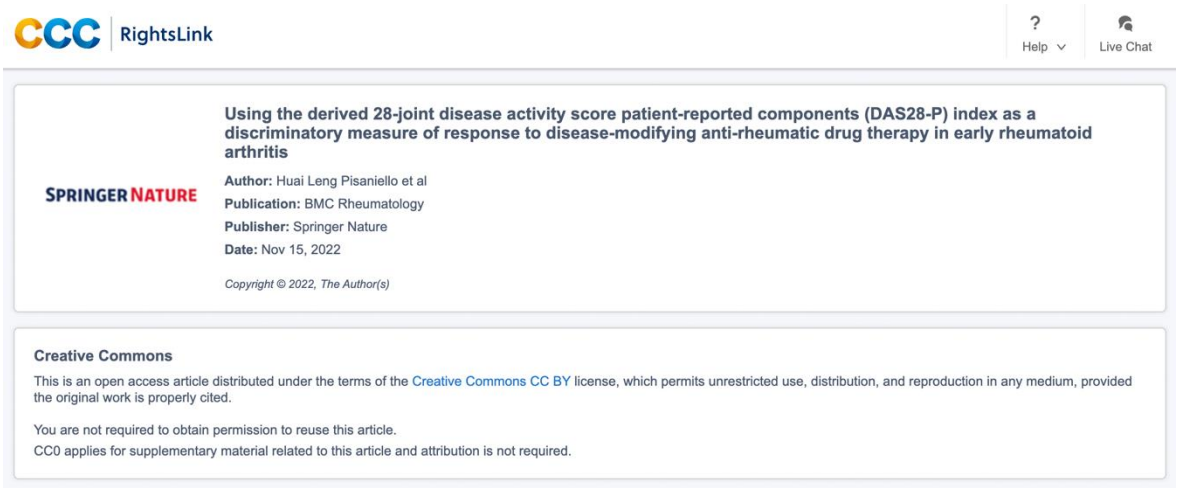
Name of Co-Author	Leah McWilliams
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 15, 2023

Name of Co-Author	Catherine L. Hill
Contribution to the Paper	Contributed to the study design and conception. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 18/04/2023

Name of Co-Author	Susanna Proudman
Contribution to the Paper	Contributed to the study design and conception, and data acquisition and management. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Jun 15, 2023

5.4 Manuscript: Using the Derived 28-Joint Disease Activity Score Patient-Reported Components (DAS28-P) Index as a Discriminatory Measure of Response to Disease Modifying Anti-Rheumatic Drug Therapy in Early Rheumatoid Arthritis

Published in *BMC Rheumatology*



BMC Rheumatology

RESEARCH

Open Access

Using the derived 28-joint disease activity score patient-reported components (DAS28-P) index as a discriminatory measure of response to disease-modifying anti-rheumatic drug therapy in early rheumatoid arthritis

Huai Leng Pisaniello[1,2*], Samuel L. Whittle[2,3], Susan Lester[2,3], Fiona Menz[4], Robert Metcalf[5], Leah McWilliams[5], Catherine L. Hill[1,2,3,5] and Susanna Proudman[1,5]

Abstract

Background: The 28-joint disease activity score (DAS28) is a widely used measure to assess disease activity in rheumatoid arthritis (RA). The DAS28-P index, a derived proportion of the patient-reported components (joint tenderness and patient global assessment) within the DAS28, has been utilized as a discriminatory measure of non-inflammatory pain mechanisms in RA. This study aimed to evaluate the use of the DAS28-P index as a predictor of treatment response in early RA.

Methods: Patients with early RA enrolled in a supplemental fish oil clinical trial received a combination of disease-modifying anti-rheumatic drugs (DMARDs) according to a 'treat-to-target' protocol. First, consecutive measures of the DAS28-P index, derived from the DAS28-erythrocyte sedimentation rate (DAS28-ESR), at each visit over a 1-year period were estimated for each patient. Then, distinct subgroups of treatment responders based on the trajectories of the DAS28-P indices were identified using bivariate k-means cluster analysis. Data on baseline predictors as well as longitudinal outcomes of disease impact and DMARD use over a 1-year period and radiographic progression over a 3-year period were collected and analyzed using a random intercept, population-averaged generalized estimating equation model.

Results: 121 patients were included (74% female; mean age of 57; median of 16 weeks of active disease) and a 3-cluster model was identified—the 'Responders' group (n = 58; 48%), the 'Partial Responders' group (n = 32; 26%), and the 'Non-Responders' group (n = 31; 26%). The 'Partial Responders' group had consistently higher proportions of the DAS28-P index throughout the study period and had minimal radiographic progression over time, with the lowest joint erosion score of 0.9 [95% confidence interval (CI) 0.2, 1.6], observed at the 3-year follow-up. At 52 weeks, the methotrexate dose was higher for both 'Partial Responders' and 'Non-Responders' groups (18.5 mg [95% CI 15.5, 21.5] and 18.6 mg [95% CI 15.3, 21.8] respectively), when compared with the 'Responders' group (12.8 mg [95% CI 14.7, 20.9]).

*Correspondence: huaileng.pisaniello@adelaide.edu.au

[2] Rheumatology Research Group, Basil Hetzel Institute for Translational Health Research, The Queen Elizabeth Hospital, Woodville South, SA, Australia
Full list of author information is available at the end of the article

Conclusions: Persistently high DAS28-P index scores are useful to distinguish poor patient global assessment and excessive treatment escalation in early RA, suggestive of underlying non-inflammatory pain contributing to higher disease activity score. Early identification of patients with discordant subjective and objective components of composite disease activity measures may allow better tailoring of treatment in RA.

Keywords: DAS28-P index, Rheumatoid arthritis, Pain, Patient global assessment

Background

Rheumatoid arthritis (RA) is a chronic autoimmune inflammatory disease that results in joint pain and swelling, as well as other peri-articular and extra-articular systemic manifestations [1]. RA characteristically affects the small joints of the hands (the metacarpophalangeal joints and the proximal interphalangeal joints), wrists, knees and feet [1]. In recent years, the long-term outcomes for RA have improved significantly, largely driven by advances in disease-modifying anti-rheumatic drugs (DMARDs), in particular biologic therapy, and with the implementation of 'early diagnosis' and 'treat-to-target' (T2T) approaches [1–4]. Pain remains a cardinal feature in RA and historically, pain mechanisms in RA have been attributed solely to activation of the peripheral nociceptive pathways by the underlying joint inflammation [5]. Consequently, the notions of controlling the disease activity with DMARDs and achieving disease remission have always been the cornerstone of pain management in RA, although this approach does not hold true for some patients with persistent pain.

In fact, the pain mechanism in RA involves a complex interaction between the inflammatory process in the joints and a combination of both the activation of the peripheral nociceptors and the peripheral and central modulation of nociceptive and other inputs [5]. The presence of joint pain despite apparent good control of synovial inflammation implies that mechanisms other than pure nociception are important in the overall pain experience in RA. The relative contributions of different peripheral and central mechanisms may vary between individuals with RA. For instance, a Swedish population-based cohort study found that nearly one-third of patients with early RA had persistent pain despite effective control of the joint inflammation. This finding was strongly predicted by having both functional impairment and low C-reactive protein (CRP) level at baseline [6]. Additionally, over-estimation of disease activity scores by non-inflammatory pain has been observed as a common occurrence in RA, irrespective of the timing of initiating or escalating treatment [7]. Identifying non-nociceptive pain during the treatment course for RA is crucial as overtreatment in patients with persistent pain unrelated to the underlying inflammation is unfavorable and can be harmful and is especially likely to occur in the context of a T2T approach.

In clinical practice and clinical trials, particularly in the modern T2T approach, the monitoring of disease activity and treatment response in RA is commonly performed by using the disease activity score 28-joints (DAS28) [8, 9]. DAS28 is a composite score derived collectively from the objective measures as assessed by the clinician (i.e., swollen joint counts (SJCs) and the acute-phase response (erythrocyte sedimentation rate (ESR) or CRP level) and the patient-reported measures (i.e., tender joint counts (TJCs) and global health as assessed by using a visual analogue scale (VAS) of patient-reported disease activity (VAS-GH)) [10, 11]. These patient-reported measures within the DAS28 composite score, representative of the patient global assessment (PGA), are more susceptible to individual-level variation due to factors other than inflammation alone. When it comes to interpreting the DAS28 score, careful interpretation of the PGA is important, especially in patients with pain driven predominantly by centrally augmented pain mechanisms [12, 13]. Various composite disease activity measures are used in the T2T paradigm. DAS28 remains in common clinical use although other measures such as the Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI) have been the preferred composite disease activity measures in recent years [14–16]. Non-inflammatory pain experienced by patients in RA may not be well captured by the overall scoring of the current composite disease activity measures, and the decomposition of a composite measure into subjective and objective components may be relevant to other measures, such as the SDAI, although this remains an open research question. McWilliams and colleague proposed that the use of the DAS28-P index, defined as "a derived measure of the proportion for the contribution of the patient-reported outcomes (TJC and VAS-GH) to the total DAS28 score", is considered a useful discriminatory index of non-inflammatory pain mechanisms in RA [12, 17]. A higher DAS28-P index was shown to predict less pain improvement in an early RA cohort at 12 months, when adjusted for baseline pain scores [12]. Similarly, a higher DAS28-P index was correlated with widespread pressure-induced pain sensitivity in established RA and with the fibromyalgia survey score [18].

In this study, building on existing research utilizing the DAS28-P index, we hypothesized that in an early RA cohort with DMARDs initiated and modified to meet a pre-defined level of disease activity, the DAS28-P index is useful in discriminating non-inflammatory pain mechanisms in early RA. We first aimed to identify different disease trajectories for each participant in this early RA cohort by using the objective and subjective components of the DAS28-ESR. Next, we aimed to assess the impact of using the subjective components of the DAS28-ESR, and therefore, the role of the DAS28-P index in monitoring disease activity and in determining the trajectory of DMARD use in a T2T approach.

Methods

Participants

Our study included a subset of participants recruited for a randomized controlled trial (RCT) of fish oil use in early RA. DAS28 was the most widely used composite disease activity measure in RA during the conduct of the study. As described in the original study, consecutive patients aged 18 years and older with early onset RA (defined as symptomatic polyarthritis of less than 12 months, SJC $\geq$ 3, TJC $\geq$ 6, ESR > 28 mm/h, and/or CRP > 10 mg/dL) diagnosed at the Rheumatology Clinic, Royal Adelaide Hospital (RAH), South Australia were recruited. These DMARD-naïve patients, who fulfilled the diagnosis of RA according to the 1987 revised American College of Rheumatology (ACR) criteria, were enrolled, and screened for eligibility to enter a double-blind RCT of high dose fish oil versus low dose fish oil. The exclusion criteria included DMARD use other than anti-malarials, use of anti-malarials for more than one month, recent seroconversion to parvovirus, Ross River, Barmah Forest, or rubella viruses, history of positive anti-nuclear antibody with a titre of $\geq$ 1:320, history of positive hepatitis B, hepatitis C or human immunodeficiency virus (HIV), known sensitivity to methotrexate, sulfasalazine or hydroxychloroquine, and history of systemic disease likely to increase risk of toxicity to 1 or more of these DMARDs. The study was approved by the RAH Research Ethics Committee (Research Protocol No: 981105).

Study protocol

The full details of the original study cohort, study design, study treatment strategy and results have been previously published elsewhere [19–21]. Alongside the randomization of receiving high dose vs low dose fish oil in the study, patients commenced DMARDs with dose adjustments based on a T2T treatment approach. In brief, triple DMARD therapy comprised methotrexate 10 mg orally weekly, folic acid 500mcg daily, sulfasalazine 500 mg daily, with dose increment over 4 weeks to 1 g twice daily and hydroxychloroquine 200 mg twice daily. The methotrexate dose was up-titrated to a maximum dose of 25 mg weekly administered subcutaneously to achieve disease remission based on pre-specified disease activity criteria [19, 20]. The addition of leflunomide was considered when maximal tolerated doses of triple DMARD therapy were achieved. Use of oral glucocorticoids and non-steroidal anti-inflammatory drugs (NSAIDs) was actively discouraged during the study, and if commenced at study inception, doses were gradually tapered and ceased where possible. Parenteral glucocorticoid use during the study was allowed as clinically indicated.

Data collections and measurements

Patients were reviewed every 3–6 weeks and measures of disease activity were taken at each follow-up visit. Specific to the aims of our study, we obtained data pertaining to longitudinal measures of disease activity and disease impact for the first 52 weeks. Both objective components (SJCs (28 joints) and ESR) and subjective components (TJCs (28 joints) and VAS-GH) of the DAS28-ESR captured for all visits (baseline and follow-up) for the first 52 weeks were reviewed and analyzed separately. On average, there were 9.7 visits per patient. A 100-mm VAS was used to assess each of these domains—global health (VAS-GH), pain and fatigue. DAS28-P index, defined as the fraction of the total DAS28-ESR contributed by the subjective components of the DAS28-ESR, was calculated for each visit. The modified Health Assessment Questionnaire (mHAQ) and 36-Item Short Form Survey (SF-36) were used at each visit to assess function and yearly to assess quality of life, respectively. A validated 5-item Rheumatology Attitudes Index (VALI-RAI) helplessness subscales (5–30) scoring system was used at each visit to evaluate patients' views of helplessness in coping with their arthritis [22]. X-rays of the hands and feet were performed for each patient annually for 3 years. In a blinded, chronological fashion, for each time point out to 3 years, two independent observers assessed the presence of joint erosion in these radiographs of hands and feet using the modified Sharp/van der Heijde (SHS) method [21, 23].

Statistical analysis

For the descriptive data, the categorical variables were summarized as absolute numbers and proportions in percentages, whereas the continuous variables were summarized as means with standard deviation or median with interquartile range. For the overall pairwise comparisons between the identified clusters, p-values for normally distributed continuous variables were calculated using one-way ANOVA and p-values for non-normally distributed continuous variables were analyzed using Kruskal–Wallis one-way ANOVA. Categorical variables were analyzed

using Pearson's chi square and 2-sided Fisher's exact tests for concomitant rheumatological diseases, such as fibromyalgia.

To identify subgroups of patients with different profiles of DAS28 trajectories, a bivariate longitudinal k-means clustering analysis was performed using both the individual objective and subjective DAS28-ESR component scores. This non-parametric clustering analysis was performed using the R package, klm3d [24]. Three treatment responder subgroups were selected according to our prior hypothesis that RA patients with persistent pain would be differentiated from both good- and non-responders.

For the identified cluster subgroups, the longitudinal outcome measures were analyzed by using a random intercept, population-averaged generalized estimating equation (GEE) model (a longitudinal generalized linear model), with an exchangeable correlation matrix and robust standard errors. In this GEE model, different regression analyses were applied for different outcome measures. For instance, binomial regression was used for both leflunomide and NSAIDs use and the presence of depression, Poisson regression for methotrexate use and negative binomial regression was used for the total joint erosion scores. The remaining outcome measures were analyzed by Gaussian regression, except the DAS28-P index scores (range of 0–1), which were analyzed by a probit fractional regression, with standard errors clustered by each patient. The outcomes for the total joint erosion scores were measured at multiple, irregularly spaced visits over the 12 months of follow-up, with an average of 9.8 months per patient. Therefore, restricted, orthogonal cubic splines (with 3 degrees of freedom (d.f.) for the main effect, and an additional 2 d.f. for interaction effects) were applied to model the responses over time. As a result, the differences between the subgroups over time were assessed by joint Wald tests of the appropriate regression coefficients.

For the SF-36 data, the scores for each nine domains were converted to Physical Component Scores and Mental Component Scores using the Stata sf36.ado module [25]. To allow for between-domain comparisons of results, each SF-36 domain scale (0–100) was transformed to a norm-based scale with a mean of 50 and standard deviation of 10, using direct age- and gender-standardization to the South Australian (SA) population norms from the 1995 National Health Survey [26].

All results were interpreted as predicted marginal means (i.e., on the original response scale) with linear contrasts used to assess differences between response from each subgroup at specific time points. The clustering analysis was performed using R version 3.2.0 and the remaining statistical analyses were performed using Stata v16.1 (StataCorp LLC, TS, USA).

Results

A total of 121 patients were included in the final analysis for this study, with 1220 observations captured from baseline to 52 weeks. These patients were predominantly female (74%) with a mean age of 57 years and had a median of 16 weeks of symptomatic polyarthritis at baseline and 54% were seropositive for anti-cyclic citrullinated peptide antibodies (ACPAs). Among these 121 patients, the k-means clustering analysis generated 3 subgroups of patients according to the 52-week trajectories of these 3 outcome measures: the overall DAS28-ESR, the objective components of the DAS28 and the subjective components of the DAS28. These subgroups were classified as Group 1—'Responders', Group 2—'Partial Responders' and Group 3—'Non-Responders'. The baseline characteristics for each of these groups are outlined in Table 1.

The predicted marginal means and their corresponding 95% confidence intervals (CI) for all outcome measures of disease activity and disease impact at baseline and at week 52 for each responder group are summarized in Table 2. Overall, at baseline, the participants in the 'Non-Responders' group were older, had a higher BMI and were more likely to have smoked (Table 1). The 'Responders' group had the lowest baseline DAS28-ESR and the 'Partial Responders' group had the highest baseline DAS28-P index (Table 2).

For the overall disease activity trajectories during the first year of treatment, the baseline mean DAS28-ESR for the whole cohort was 5.7 (s.d. 1.2). At 52 weeks, the DAS28-ESR score in the 'Partial Responders' group was 3.9 [95% CI 3.3, 4.4], which was worse than the 'Responders' group (2.3 [95% CI 2.1, 2.6; $p < 0.05$]) and better than the 'Non-Responders' group (5.1 [95% CI 4.7, 5.5; $p < 0.05$]). These results were largely due to the lower subjective DAS28 score of 0.6 [95% CI 0.5, 0.8; $p < 0.05$] for the 'Responders' group and higher objective DAS28 score of 2.8 [95% CI 2.7, 3.0; $p < 0.05$] for the 'Non-Responders' group. As shown in Fig. 1, the overall DAS28-P mean scores during the first year of treatment were consistently above 0.5 for the 'Partial Responder' group in comparison with the 'Responders' and 'Non-Responders' groups, with the subjective DAS28 score being the major contributor to the total DAS28 score. Notably, the trajectory of the objective DAS28 scores in the 'Partial Responder' group was similar to the 'Responder' group (Fig. 1B), whereas the trajectory of the subjective DAS28 scores in the 'Partial Responder' group was similar to the 'Non-Responder' group (Fig. 1C).

Table 1 Baseline characteristics of participants included in the study

Descriptor	Group 1	Group 2	Group 3	All
Classification	Responders	Partial responders	Non-responders	
N (%)	58 (48%)	32 (26%)	31 (26%)	121
Age: mean (SD)	56 (16)	54 (15)	63 (11)	57 (15)
Females (%)	45 (78%)	25 (78%)	19 (61%)	89 (74%)
BMI: mean (SD)	26.4 (5.1)	27.4 (4.8)	30.6 (7.6)*	27.7 (6.0)
ACPA positive (%)	31/55 (56%)	14/31 (45%)	18/31 (58%)	63/117 (54%)
RF positive (%)	39/55 (71%)	13/31 (42%)	18/31(58%)	70/117 (60%)
Smoking (%)				
Never	29 (50%)	16 (50%)	7 (23%)	52 (43%)
Former	20 (34%)	14 (44%)	15 (48%)	49 (41%)
Current	9 (16%)	2 (6%)	9 (29%)	20 (17%)
Odds ratio$_{ordinal}$ (95% CI) for the likelihood of smoking	*1.2 (0.5, 2.7)*	*1*	*3.6 (1.4, 9.2)**	
Weeks polyarthritis: median (IQR)	16 (12, 24)	16 (12, 24)	20 (12, 28)*	16 (12, 24)
Randomized to fish oil (%)	38 (66%)	20 (63%)	17 (54%)	75 (62%)

The odds ratio (highlighted in italics) was calculated for the likelihood of smoking

* Significantly different from Group 2 ($p < 0.05$)

SD standard deviation, *BMI* body mass index, *ACPA* anti-cyclic citrullinated peptide antibody, *RF* rheumatoid factor, *CI* confidence interval, *IQR* interquartile range

Table 2 Predicted marginal means (95% confidence intervals, CI), by responder group, for all outcome measures at baseline and at week 52

	Baseline			52 Weeks		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
	Responders	Partial responders	Non-responders	Responders	Partial responders	Non-responders
DAS28-ESR	5.1 (4.9, 5.4)*	6.1 (5.7, 6.5)	6.3 (5.9, 6.6)	2.3 (2.1, 2.6)*	3.9 (3.3, 4.4)	5.1 (4.7, 5.5)*
DAS28-ESR objective component	3.0 (2.9, 3.2)	3.0 (2.7, 3.2)	3.4 (3.2, 3.6)*	1.7 (1.5, 1.8)	1.7 (1.5, 1.9)	2.8 (2.7, 3.0)*
DAS28-ESR subjective component	2.1 (1.9, 2.3)*	3.1 (2.9, 3.4)	2.9 (2.7, 3.1)	0.6 (0.5, 0.8)*	2.1 (1.8, 2.5)	2.3 (2.0, 2.6)
DAS28-P	0.40 (0.38, 0.43)*	0.52 (0.50, 0.55)	0.45 (0.43, 0.48)*	0.23 (0.19, 0.28)*	0.56 (0.52, 0.59)	0.43 (0.39, 0.47)*
mHAQ	0.61 (0.49, 0.73)*	0.91 (0.75, 1.07)	0.82 (0.64, 0.99)	0.11 (0.04, 0.18)*	0.47 (0.25, 0.67)	0.48 (0.32, 0.63)
Fatigue	36.1(30.6, 41.6)*	66.7 (58.7, 74.6)	53.7 (45.1, 62.2)*	18.9 (12.7, 25.1)*	44.6 (35.0, 54.1)	48.3 (39.4, 57.1)
Helplessness	13.0 (11.8, 14.3)*	16.1 (14.2, 17.9)	15.4 (13.7, 17.1)	8.6 (7.6, 9.6)*	13.0 (11.4, 14.6)	13.0 (11.4, 14.7)
SF-36						
Physical component score	37.5 (35.1, 39.9)*	32.1 (29.9, 34.3)	33.5 (30.6, 36.4)	47.6 (45.3, 50.0)*	38.5 (35.4, 41.6)	35.7 (32.9, 38.6)
Mental component score	43.5 (40.8, 46.3)*	35.3 (33.3, 38.3)	37.0 (32.8, 41.2)	48.4 (40.8, 46.3)*	41.4 (37.6, 45.1)	40.1 (35.9, 44.4)
Methotrexate dose				12.8 (14.7, 20.9)*	18.5 (15.5, 21.5)	18.6 (15.3, 21.8)
Leflunomide use (%)				4% (0, 10)*	16% (2, 30)	40% (21, 59)
Cumulative glucocorticoid dose, mg: median (IQR)				171 (100, 250)	199 (150, 450)	297 (211, 484)
Total erosion score	0.5 (0.3, 0.8)	0.5 (0.1, 1.0)	1.0 (0.3, 1.7)	1.2 (0.6, 1.7)	0.9 (0.2, 1.7)	2.4 (1.1, 3.6)*

The cumulative glucocorticoid dose was calculated in milligrams(mg) of prednisolone equivalent

DAS28-ESR disease activity score 28-joints-Erythrocyte Sedimentation rate, *mHAQ* modified Health Assessment Questionnaire, *SF-36* 36-item short form survey, *IQR* interquartile range

* Significantly different from Group 2 ($p < 0.05$)

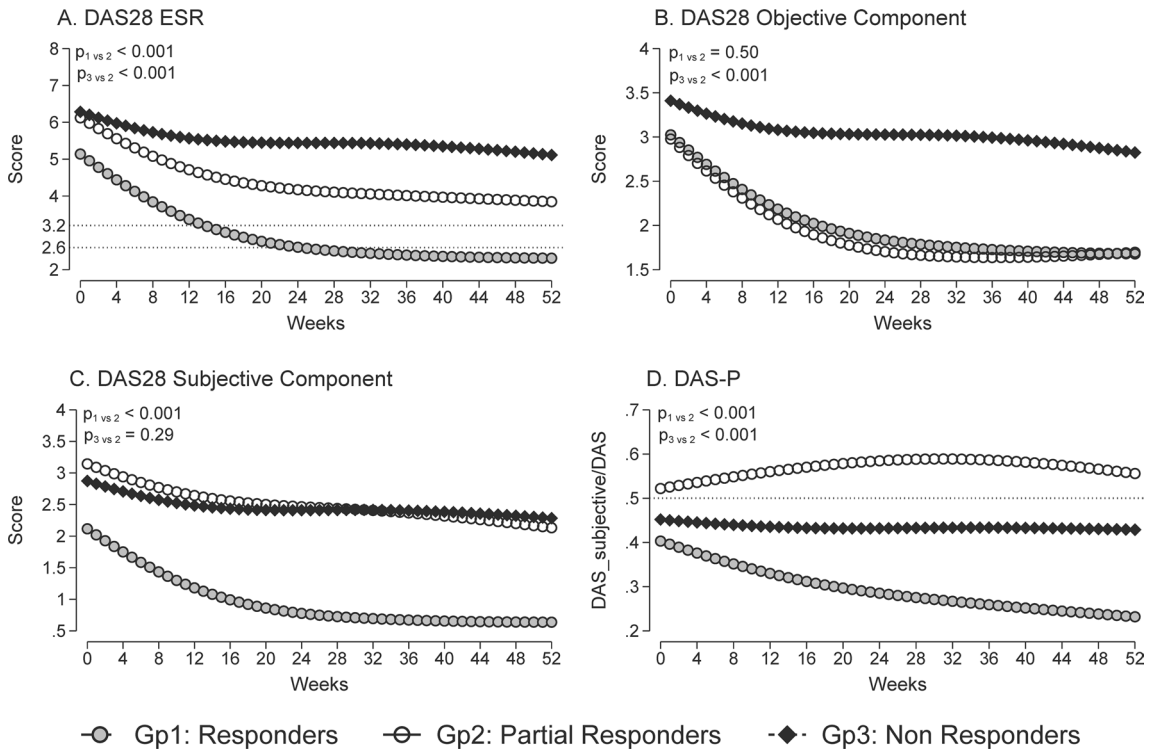


Fig. 1 DAS28 and DAS28 component scores (predicted marginal means) during the first year of treatment for the three responder groups. **A** DAS28-ESR **B** DAS28—Objective Component **C** DAS28—Subjective Component **D** DAS28-P (an index defined by DAS28—Subjective Component/DAS28-ESR)

In terms of self-reported outcome measures of disease impact during the first year of treatment, the overall mean mHAQ at baseline for the whole cohort was 0.75 (s.d. 0.54). As shown in Fig. 2, the 'Responders' group had consistently lower mHAQ scores when compared to the 'Partial Responders' group ($p < 0.001$). Similarly, the 'Responders' group had consistently lower levels of fatigue and helplessness scores ($p < 0.001$ for both measures) when compared to the 'Partial Responders' group. Both the physical and mental component scores of the SF-36 were statistically better in the 'Responders' group (means of 47.6 [95% CI 45.3, 50.0] and 48.4 [95% CI 40.8, 46.3] respectively).

In terms of DMARD treatment comparisons between the responder groups during the first year of treatment, based on the T2T approach, the methotrexate dose profile was higher in the 'Partial Responders' group compared to the 'Responders' group, but was comparable to the 'Non-Responders' group (Fig. 3A). At 52 weeks, the methotrexate dose (in milligram, mg) was 12.8 mg [95% CI 14.7, 20.9] for the 'Responders' group, 18.5 mg [95% CI 15.5, 21.5] for the 'Partial Responders' group and 18.6 mg [95% CI 15.3, 21.8] for the 'Non-Responders' group. The leflunomide use profile in the 'Partial Responders' group was intermediate between the 'Responders' and 'Non-Responders' groups (Fig. 3B). Although NSAID use was permitted during the study period, there were no significant differences in terms of NSAID use profile between the three responder groups. In patients who were given glucocorticoid during the study period, the cumulative glucocorticoid doses (expressed in milligrams of prednisolone equivalent) at week 52 were relatively low among the three responder groups ($p = 0.022$, Kruskal–Wallis rank test for equality of populations). The median cumulative glucocorticoid dose was the highest in the 'Non-Responders' group (297 mg, IQR 211, 284; $n = 24$), compared to the 'Partial Responders' group (199 mg, IQR 150, 450; $n = 24$) and the 'Responders' group (171 mg, IQR 100, 250; $n = 27$) (Table 2).

The total joint erosion scores at baseline and 12 months follow-up are outlined in Table 2. While the comparison of total joint erosion scores between the 'Responders' and 'Partial Responders' groups over all time points was significant ($p = 0.033$), there were no significant differences at any individual time point (Fig. 4). However, the comparison between the 'Partial Responders' and 'Non-Responders' groups was significant ($p < 0.05$) at each follow-up time point (years 1, 2 and 3). At year 3, the total

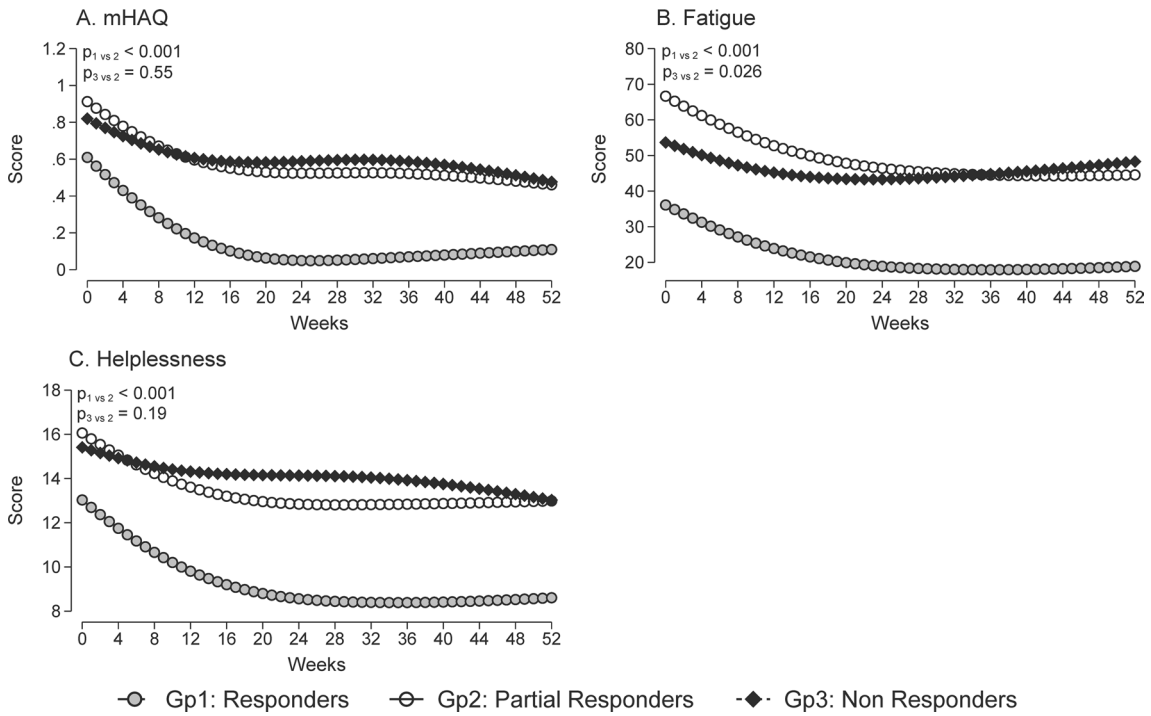


Fig. 2 Self-reported scores (predicted marginal means) during the first year of treatment for the three responder groups: **A** modified Health Assessment Questionnaire (mHAQ) disability **B** Fatigue (measured in 0–100 mm visual analogue scale) **C** Helplessness in coping with arthritis (measured in a 5–30 subscale scoring system in a Validated 5-item Rheumatology Attitudes Index (VALI-RAI))

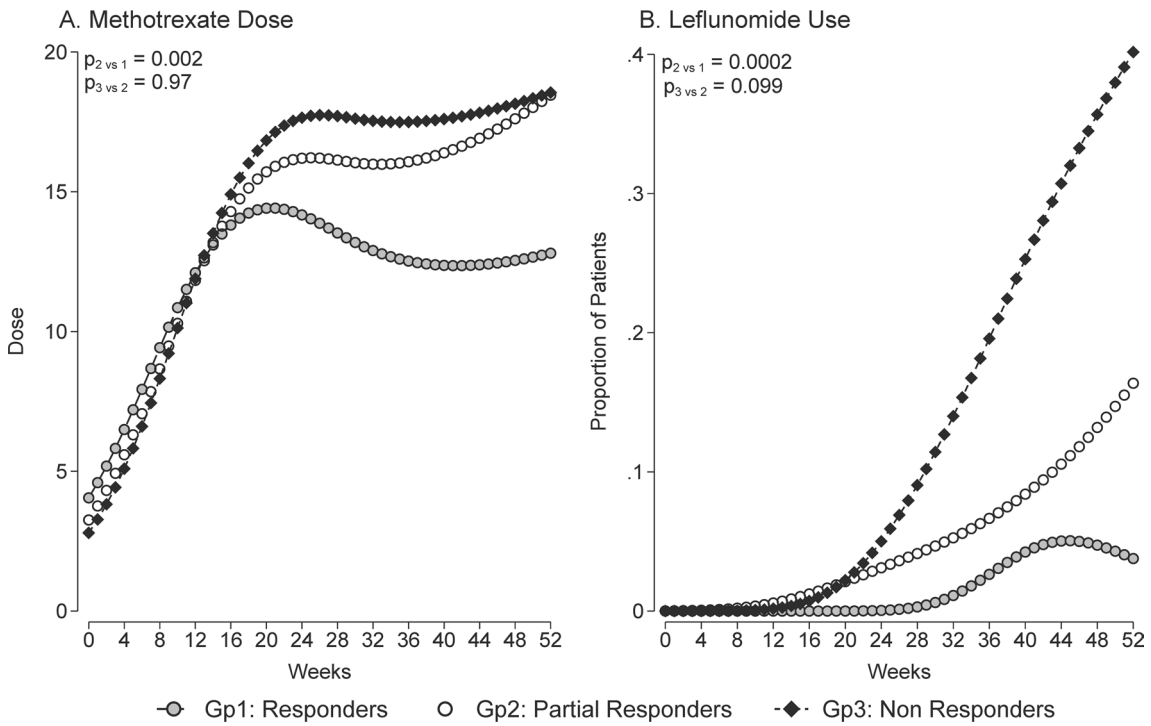


Fig. 3 DMARD treatment (predicted marginal means) during the first year for the three responder groups: **A** Methotrexate (MTX) Dose (in milligram, mg of weekly dosing) **B** Leflunomide Use (proportion of patients)

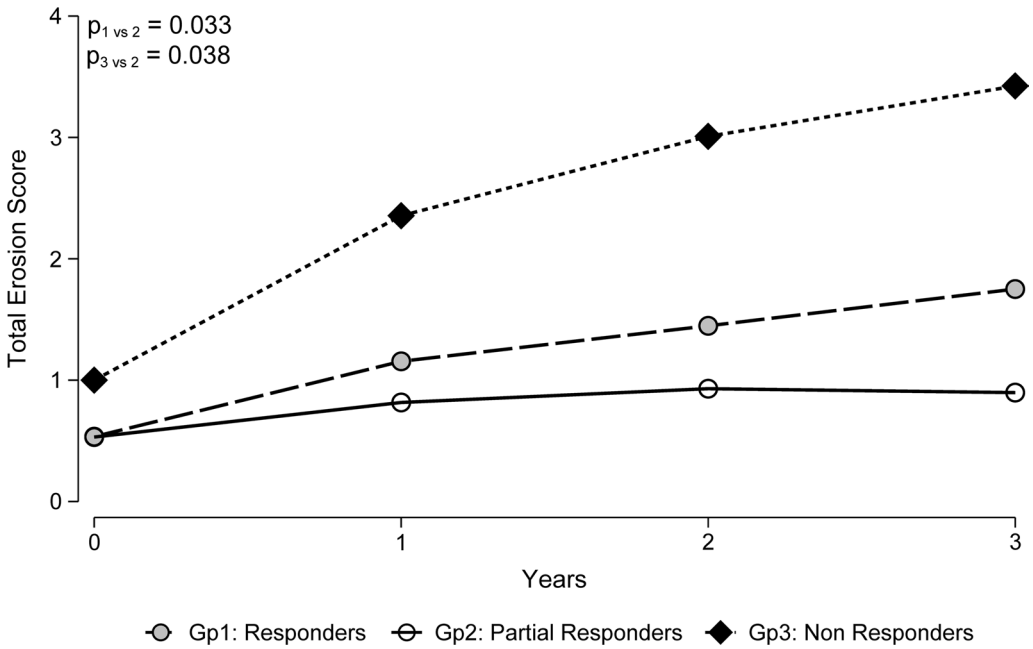


Fig. 4 Total Erosion Scores (predicted marginal means) over 3 years of follow-up for the three responder groups

joint erosion score in the 'Responders' group was 1.8 (95% CI 0.8, 2.7), compared to 0.9 (95% CI 0.2, 1.6) in the 'Partial Responders' group and 3.4 (95% CI 1.7, 5.2) in the 'Non-Responders' group.

Discussion

Composite disease activity measures such as the DAS28 are routinely used in rheumatology practice to monitor the disease trajectory in RA. In this study of an early RA cohort managed with a T2T approach, we identified three distinct subgroups of patients with different disease trajectories over 12 months by clustering each component of the DAS28 (the overall score, and both the objective and subjective components of the score). Of the 121 study participants, at 52 weeks, nearly half of them were in disease remission (the 'Responders' group), and the other half of the study cohort continued to have moderate-to-high disease activity, with 26% in the 'Partial Responders' group and 26% in the 'Non-Responders' group. When we examined the components of the DAS28, both 'Responders' and 'Partial Responders' groups had similar DAS28 objective component mean scores at baseline and at 52 weeks, and yet, these two subgroups had different disease trajectories. In fact, the relatively high total mean DAS28-ESR scores for the 'Partial Responders' group at baseline and at 52 weeks were largely driven by the reporting of high DAS28 subjective component scores, as highlighted by the consistently higher proportions of the DAS28-P index throughout the study period when compared to both 'Responders' and 'Non-Responders' groups. Despite receiving similar T2T therapy, these findings in the 'Partial Responders' group reflect ongoing patient-reported concerns about their disease trajectories disproportionate to the underlying disease inflammation.

In our study, both 'Partial Responders' and 'Non-Responders' groups reported similar worsening of disease impact throughout the study period, when compared to the 'Responders' group, as demonstrated in the mHAQ scores, the level of fatigue and helplessness scores. Apart from the higher level of fatigue in the 'Partial Responders' group, both 'Partial Responders' and 'Non-Responders' groups were indistinguishable at baseline, even in the DAS28 subjective component scores. Evidently, these two subgroups differed in the trajectories of the DAS28 objective component scores and the DAS28-P proportion indices. Again, according to the DAS28-P index, these findings suggest a predominance of non-inflammatory pain mechanisms in the 'Partial Responders' group and failure of treatment and ongoing active disease in the 'Non-Responders' group at baseline and throughout the study. Although a difference in the DAS28-P was seen between the 'Partial Responders' group and the 'Non-Responders' group (Fig. 1D), the difference between each group may not be sufficient to reliably categorize individual patients at any time point. Other concomitant chronic pain conditions, such as osteoarthritis and fibromyalgia, could be confounders for persistent pain

in this study cohort, and although beyond the scope of our study, their contribution to non-inflammatory pain in early RA would be relevant in any future analysis. Although the DAS28-P index can be used as a discriminatory measure of non-inflammatory pain in RA, our study highlights that baseline DAS28-P does not predict trajectory of RA disease activity in individuals, which was not previously examined in the original study proposing the use of DAS28-P index [12].

Overtreatment is a potential risk in the modern treatment era for patients diagnosed with early RA, especially in the T2T approach [27]. In this study, there was a substantial increase in both the methotrexate mean dose and the proportions of leflunomide users in both 'Partial Responders' and 'Non-Responders' groups. In detail, dose increments for both of these DMARDs were seen at week 16, a typical time period for deciding any change in dosing, and subsequently the doses were gradually uptitrated to the maximum recommended target doses, as dictated by the serial DAS28 scores. Similarly, despite the analysis of only a subset of the study cohort, the cumulative dose of glucocorticoid use in the 'Non-Responders' group was substantively higher compared to the other two subgroups. Consequently, these subgroups with disproportionate dose titration and disease activity could be at risk of DMARD-related toxicity in the intermediate- and long-term. Likewise, a recent study by Wallace and colleague revealed two thirds of established rheumatoid arthritis patients had persistent glucocorticoid use, especially in those with high fibromyalgianess [28]. In our study, we observed that a higher DAS28-P in both 'Partial Responders' and 'Non-Responders' groups was associated with higher exposure to combination DMARD therapy. Relying on the use of only the composite DAS28 score might lead to overtreatment, which could be mitigated by understanding the relative contributions of subjective and objective measures to the total composite score. In addition, escalation to biologic DMARDs in these subgroups may occur, which may result in higher societal and health care cost and unnecessary immunosuppression. Future studies examining the use of conventional synthetic DMARDs beyond a 1-year period and the timing of switching to biologic DMARD in these subgroups of early RA cohort may help to further characterize the impact of the T2T treatment approach in those with persistent non-inflammatory pain in RA.

Furthermore, in our study, despite not achieving the DAS28 indicative of low disease activity/disease remission, the 'Partial Responders' group had the lowest joint erosion scores serially over 3 years, demonstrating no progression of erosive disease. This is consistent with their low levels of disease inflammation following treatment, as reflected by the overall DAS28 objective component score. This finding underscores the risk of unnecessary overtreatment in a 'partially-responsive' subgroup of patients with early RA, in whom additional immunosuppressive agents will not alleviate non-inflammatory symptoms. Adjuvant interventions that target non-inflammatory pain rather than relying on immunosuppressive therapies are likely warranted in this group of patients with suboptimal disease control despite no objective evidence of ongoing inflammation [29, 30]. With regards to radiographic progression, the 'Responders' group had higher erosion risks compared to the 'Partial Responders' group, although the 'Responders' group had the overall lowest subjective DAS28-ESR scores. This may reflect the recognised phenomenon of progressive structural damage even when objective measures of disease activity are low/normal, and highlights the importance of assessing radiographic outcomes in addition to both subjective and objective disease activity measures in RA [31–33]. Judicious interpretation of all these outcome measures may lower the risk of overtreatment in those with high DAS28-P and, conversely, undertreatment in those with low DAS28-P.

In the modern T2T strategies in achieving disease remission in RA, we are yet to have mutually exclusive composite measures to incorporate disease outcome measures important to both clinicians and patients. From the patient's perspective, disease remission comprises both resolution of disease inflammation and alleviation of symptoms related to the disease. Although the PGA within the DAS28 has been considered the cornerstone of determining the patient-reported disease remission, the role of PGA remains contentious. A recent large individual patient data meta-analysis evaluating the impact of PGA in the definition of disease remission and as a predictor of radiographic damage in RA concluded that the current DAS28 remission definition that includes the PGA, is better than a definition that excludes PGA for predicting a good functional outcome but reduces the predictive accuracy for radiological outcomes, raising concerns for risk of overtreatment [34]. In a large multinational study using the METEOR database of patients on biologic DMARDs for RA, the PGA remained high in those in remission, with the danger of further unnecessary immunosuppression [35]. Ferreira and colleagues have proposed a dual T2T strategy, which comprises the management of disease inflammation (biologic remission) and the management of disease impact (symptom remission) to guide treatment in RA [36]. For biologic remission, alongside the dual T2T, the author recommended the use of 3-variable remission—SJC, TJC and CRP [37]. For symptom remission, the author suggested further validation of the PGA with the use of the Rheumatoid Arthritis Impact of Disease (RAID) score [37].

This additional patient-reported measure in early RA may provide early insight at the start of DMARD initiation, with early adjuvant interventions to be provided to those who are likely to have persistent non-inflammatory pain.

Our study has some strengths and limitations. Our study examined patients who were definitively diagnosed with early RA as they were recruited through strict RCT inclusion criteria. We were able to differentiate patients with persistent pain (the 'Partial Responders' group) from the 'Non-Responders' group, a difference that was not shown in the previous study using the DAS28-P index [12]. It included a relatively small cohort of patients with early RA. Irrespective, we had adequate study size and repeated measures to provide representative subgroups of patients with different trajectories to evaluate the utility of DAS28-P index as predictor of treatment response in the first year after diagnosis of RA. In addition, the study participants were mainly recruited from a single-center rheumatology unit, which may introduce selection bias in terms of the residency of the patients and their corresponding education levels and socio-economic status.

In summary, in this well-characterized early RA cohort managed with a T2T approach within the first year, the DAS28-P index can be used as a discriminatory measure of non-inflammatory pain in RA, but baseline DAS28-P does not necessarily predict trajectory in individuals. Concurrent assessment of both objective and subjective components of the DAS28 is likely to be most informative when it comes to tailoring of therapy in patients with RA, especially in treatment escalation. Most importantly, early identification of patients with discordant subjective and objective outcomes may facilitate optimal shared decision-making regarding DMARD and pain management. Additional clinical assessment and communication are warranted when there is a suspicion of ongoing non-inflammatory pain despite adequate control of disease inflammation.

Abbreviations

ACR: American College of Rheumatology; ACPAs: Anti-cyclic citrullinated peptide antibodies; CI: Confidence intervals; CRP: C-reactive protein; DAS28: Disease activity score 28-joints; d.f.: Degrees of freedom; DMARDs: Disease-modifying anti-rheumatic drugs; ESR: Erythrocyte sedimentation rate; GEE: Generalized estimating equation; g: Gram; mcg: Microgram; mg: Milligram; mHAQ: Modified Health Assessment Questionnaire; NSAIDs: Non-steroidal anti-inflammatory drugs; PGA: Patient global assessment; RCT: Randomized controlled trial; RA: Rheumatoid arthritis; RAH: Royal Adelaide Hospital; RAID: Rheumatoid Arthritis Impact of Disease; s.d.: Standard deviation; SF-36: 36-Item Short Form Survey; SHS: Sharp/van der Heijde; SA: South Australian; SJCs: Swollen joint counts; TJCs: Tender joint counts; T2T: Treat-to-target; VALI-RAI: Validated 5-item Rheumatology Attitudes Index; VAS: Visual analogue scale; VAS-GH: Visual analogue scale (VAS) of patient-reported disease activity.

Acknowledgements

For the original fish oil study, Leah McWilliams and Cindy Hall were involved in the database management and the general organisation for the study participants. Several rheumatology advanced trainee doctors assisted in the early arthritis rheumatology clinic. The Royal Adelaide Hospital Pharmacy provided the organisation and preparation of the study oil.

Author contributions

HLP, SW, SL, FM, CH and SP contributed to the conception and design of the work, the acquisition, analysis and interpretation of the data. HLP contributed to the main drafting and the writing of the manuscript and all authors (HLP, SLW, SL, FM, RM, LM, CLH, SP) contributed to the subsequent drafting and revision of the manuscript. All authors read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. This manuscript formed part of HLP's Ph.D. work, of which HLP was the recipient of the Ken Muirden Overseas Training Fellowship from the Arthritis Australia, an educational grant funded by the Australian Rheumatology Association, and HLP also received the Australia Postgraduate Scholarship from the University of Adelaide.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Declarations

Ethics approval and consent to participate

The study was approved by the RAH Research Ethics Committee (Research Protocol No: 981105). All study methods were carried out in accordance with relevant guidelines and regulations. Informed consent was obtained from all study participants.

Consent for publication

Not applicable.

Competing interests

The authors declared having no competing interests in relation to this manuscript.

Author details

[1]Discipline of Medicine, Faculty of Health and Medical Sciences, The University of Adelaide, Adelaide, SA, Australia. [2]Rheumatology Research Group, Basil Hetzel Institute for Translational Health Research, The Queen Elizabeth Hospital, Woodville South, SA, Australia. [3]Rheumatology Unit, The Queen Elizabeth Hospital, Woodville South, SA, Australia. [4]Flinders Medical Centre, Bedford Park, SA, Australia. [5]Rheumatology Unit, Royal Adelaide Hospital, Adelaide, SA, Australia.

Received: 1 March 2022 Accepted: 7 July 2022
Published online: 15 November 2022

References

1. Sparks JA. Rheumatoid arthritis. Ann Intern Med. 2019;170(1):ITC1–16. https://doi.org/10.7326/AITC201901010.
2. Lacaille D, Avina-Zubieta JA, Sayre EC, Abrahamowicz M. Improvement in 5-year mortality in incident rheumatoid arthritis compared with the general population-closing the mortality gap. Ann Rheum Dis. 2017;76(6):1057–63. https://doi.org/10.1136/annrheumdis-2016-209562.
3. Markusse IM, Akdemir G, Dirven L, Goekoop-Ruiterman YP, van Groenendael JH, Han KH, et al. Long-term outcomes of patients with recent-onset rheumatoid arthritis after 10 years of tight controlled treatment: a randomized trial. Ann Intern Med. 2016;164(8):523–31. https://doi.org/10.7326/M15-0919.
4. Zhang Y, Lu N, Peloquin C, Dubreuil M, Neogi T, Avina-Zubieta JA, et al. Improved survival in rheumatoid arthritis: a general population-based

cohort study. Ann Rheum Dis. 2017;76(2):408–13. https://doi.org/10.1136/annrheumdis-2015-209058.
5. Zhang A, Lee YC. Mechanisms for joint pain in rheumatoid arthritis (RA): from cytokines to central sensitization. Curr Osteoporos Rep. 2018;16(5):603–10. https://doi.org/10.1007/s11914-018-0473-5.
6. Altawil R, Saevarsdottir S, Wedren S, Alfredsson L, Klareskog L, Lampa J. Remaining pain in early rheumatoid arthritis patients treated with methotrexate. Arthritis Care Res. 2016;68(8):1061–8. https://doi.org/10.1002/acr.22790.
7. Christensen AW, Rifbjerg-Madsen S, Christensen R, Dreyer L, Tillingsoe H, Seven S, et al. Non-nociceptive pain in rheumatoid arthritis is frequent and affects disease activity estimation: cross-sectional data from the FRAME study. Scand J Rheumatol. 2016;45(6):461–9. https://doi.org/10.3109/03009742.2016.1139174.
8. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJ, Hazes JM, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. Arthritis Rheum. 2008;58(2 Suppl):S126–35. https://doi.org/10.1002/art.23364.
9. Nikiphorou E, Norton S, Young A, Carpenter L, Dixey J, Walsh DA, et al. Association between rheumatoid arthritis disease activity, progression of functional limitation and long-term risk of orthopaedic surgery: combined analysis of two prospective cohorts supports EULAR treat to target DAS thresholds. Ann Rheum Dis. 2016;75(12):2080–6. https://doi.org/10.1136/annrheumdis-2015-208669.
10. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995;38(1):44–8. https://doi.org/10.1002/art.1780380107.
11. van der Heijde DM, van't Hof MA, van Riel PL, Van de Putte L. Disease activity score. Ann Rheum Dis. 1992;51(1):140.
12. McWilliams DF, Zhang W, Mansell JS, Kiely PD, Young A, Walsh DA. Predictors of change in bodily pain in early rheumatoid arthritis: an inception cohort study. Arthritis Care Res. 2012;64(10):1505–13. https://doi.org/10.1002/acr.21723.
13. Nikiphorou E, Radner H, Chatzidionysiou K, Desthieux C, Zabalan C, van Eijk-Hustings Y, et al. Patient global assessment in measuring disease activity in rheumatoid arthritis: a review of the literature. Arthritis Res Ther. 2016;18(1):251. https://doi.org/10.1186/s13075-016-1151-6.
14. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis. 2010;69(9):1580–8. https://doi.org/10.1136/ard.2010.138461.
15. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Ann Rheum Dis. 2011;70(3):404–13. https://doi.org/10.1136/ard.2011.149765.
16. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology. 2003;42(2):244–57. https://doi.org/10.1093/rheumatology/keg072.
17. McWilliams DF, Kiely PDW, Young A, Joharatnam N, Wilson D, Walsh DA. Interpretation of DAS28 and its components in the assessment of inflammatory and non-inflammatory aspects of rheumatoid arthritis. BMC Rheumatol. 2018;2:8. https://doi.org/10.1186/s41927-018-0016-9.
18. Joharatnam N, McWilliams DF, Wilson D, Wheeler M, Pande I, Walsh DA. A cross-sectional study of pain sensitivity, disease-activity assessment, mental health, and fibromyalgia status in rheumatoid arthritis. Arthritis Res Ther. 2015;17:11. https://doi.org/10.1186/s13075-015-0525-5.
19. Proudman SM, James MJ, Spargo LD, Metcalf RG, Sullivan TR, Rischmueller M, et al. Fish oil in recent onset rheumatoid arthritis: a randomised, double-blind controlled trial within algorithm-based drug use. Ann Rheum Dis. 2015;74(1):89–95. https://doi.org/10.1136/annrheumdis-2013-204145.
20. Proudman SM, Keen HI, Stamp LK, Lee AT, Goldblatt F, Ayres OC, et al. Response-driven combination therapy with conventional disease-modifying antirheumatic drugs can achieve high response rates in early rheumatoid arthritis with minimal glucocorticoid and nonsteroidal anti-inflammatory drug use. Semin Arthritis Rheum. 2007;37(2):99–111. https://doi.org/10.1016/j.semarthrit.2007.02.001.
21. Wechalekar MD, Lester S, Hill CL, Lee A, Rischmueller M, Smith MD, et al. Active foot synovitis in patients with rheumatoid arthritis: unstable remission status, radiographic progression, and worse functional outcomes in patients with foot synovitis in apparent remission. Arthritis Care Res. 2016;68(11):1616–23. https://doi.org/10.1002/acr.22887.
22. DeVellis RF, Callahan LF. A brief measure of helplessness in rheumatic disease: the helplessness subscale of the Rheumatology Attitudes Index. J Rheumatol. 1993;20(5):866–9.
23. van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. J Rheumatol. 1999;26(3):743–5.
24. Genolini C, Alacoque X, Sentenac M, Arnaud C. kml and kml3d: R packages to cluster longitudinal data. J Stat Softw. 2015;65(4):66.
25. Ryan P. Stata module to calculate summary statistics for the SF-36 Health Survey Instrument. 1999.
26. South Australian population norms for the Short Form (SF-36) Health Status Questionnaire: Unit BE; 1995.
27. Landewe RBM. Overdiagnosis and overtreatment in rheumatology: a little caution is in order. Ann Rheum Dis. 2018;77(10):1394–6. https://doi.org/10.1136/annrheumdis-2018-213700.
28. Wallace BI, Moore MN, Heisler AC, Muhammad LN, Song J, Clauw DJ, et al. Fibromyalgianess and glucocorticoid persistence among patients with rheumatoid arthritis. Rheumatology. 2021. https://doi.org/10.1093/rheumatology/keab583.
29. de Hair MJH, Jacobs JWG, Schoneveld JLM, van Laar JM. Difficult-to-treat rheumatoid arthritis: an area of unmet clinical need. Rheumatology. 2018;57(7):1135–44. https://doi.org/10.1093/rheumatology/kex349.
30. Santos EJF, Duarte C, Ferreira RJO, Pinto AM, Geenen R, da Silva JAP, et al. Determinants of happiness and quality of life in patients with rheumatoid arthritis: a structural equation modelling approach. Ann Rheum Dis. 2018;77(8):1118–24. https://doi.org/10.1136/annrheumdis-2017-212934.
31. Plant MJ, Williams AL, O'Sullivan MM, Lewis PA, Coles EC, Jessop JD. Relationship between time-integrated C-reactive protein levels and radiologic progression in patients with rheumatoid arthritis. Arthritis Rheum. 2000;43(7):1473–7. https://doi.org/10.1002/1529-0131(200007)43:7%3c1473::AID-ANR9%3e3.0.CO;2-N.
32. van Leeuwen MA, van Rijswijk MH, Sluiter WJ, van Riel PL, Kuper IH, van de Putte LB, et al. Individual relationship between progression of radiological damage and the acute phase response in early rheumatoid arthritis. Towards development of a decision support system. J Rheumatol. 1997;24(1):20–7.
33. van Leeuwen MA, van Rijswijk MH, van der Heijde DM, Te Meerman GJ, van Riel PL, Houtman PM, et al. The acute-phase response in relation to radiographic progression in early rheumatoid arthritis: a prospective study during the first three years of the disease. Br J Rheumatol. 1993;32(Suppl 3):9–13. https://doi.org/10.1093/rheumatology/32.suppl_3.9.
34. Ferreira RJO, Welsing PMJ, Jacobs JWG, Gossec L, Ndosi M, Machado PM, et al. Revisiting the use of remission criteria for rheumatoid arthritis by excluding patient global assessment: an individual meta-analysis of 5792 patients. Ann Rheum Dis. 2020. https://doi.org/10.1136/annrheumdis-2020-217171.
35. Ferreira RJO, Carvalho PD, Ndosi M, Duarte C, Chopra A, Murphy E, et al. Impact of patient's global assessment on achieving remission in patients with rheumatoid arthritis: a multinational study using the METEOR database. Arthritis Care Res. 2019;71(10):1317–25. https://doi.org/10.1002/acr.23866.
36. Ferreira RJO, Duarte C, Ndosi M, de Wit M, Gossec L, da Silva JAP. Suppressing inflammation in rheumatoid arthritis: does patient global assessment blur the target? A practice-based call for a paradigm change. Arthritis Care Res. 2018;70(3):369–78. https://doi.org/10.1002/acr.23284.
37. Ferreira RJO, Ndosi M, de Wit M, Santos EJF, Duarte C, Jacobs JWG, et al. Dual target strategy: a proposal to mitigate the risk of overtreatment and enhance patient satisfaction in rheumatoid arthritis. Ann Rheum Dis. 2019;78(10): e109. https://doi.org/10.1136/annrheumdis-2018-214199.

Publisher's Note

5.5 Summary

28-joint disease activity score (DAS28) is a common and widely used composite disease activity score in disease monitoring and treatment guidance in RA. This study demonstrated the success in using the DAS28-P index score as an appropriate discriminatory measure of response to DMARD therapy in a single-site early RA cohort.

Main findings include:

- Using a k-means clustering method on 1-year period data of DAS28-ESR of 121 study participants, three distinct trajectories of disease activity were identified – 'Responders' (n = 58; 48%), 'Partial Responders' (n = 32; 26%), and 'Non-Responders' (n = 31; 26%)
- Unnecessary treatment escalation was seen in the 'Partial Responders' group in the context persistently high levels of DAS28 (primarily driven by the subjective component) but had the lowest joint erosion scores, implying the precarious influence of the PGA component on treatment decision
- In contrast, the 'Responders' group had the lowest levels of the objective components of the DAS28 scores and yet had the highest joint erosion scores over time, suggesting the importance of ongoing monitoring of both radiographic progression and disease activity composite measure in RA
- Baseline DAS28-P does not predict disease trajectory in RA, although DAS28-P index has been shown to be a good discriminatory measure of non-inflammatory pain in RA
- The PGA role in DAS28 score is contentious especially in the current treat-to-target (T2T) strategy and biologic DMARD use in RA, and more concerningly, in those with predominantly non-inflammatory pain, suggesting that dual T2T approach (biologic remission and symptom remission) may be necessary

Chapter 6: Trajectories of Pain-Related Health Status in RA and the Associations with Sociodemographic Factors and Treatment

6.1 Preface

The study presented in this chapter addresses the final part of the research question of investigating the trajectories of pain symptom and pain-related health outcomes. Using the identified trajectory subgroups as predictors and capturing at-risk subgroup of persistent pain with poorer pain-related health status, this study aimed to investigate the associations with baseline demographics, sociodemographic indicators, comorbidities, and time-varying effects on medication use.

This study is presented in a manuscript format in this chapter, which has been published in a peer-reviewed journal, the RMD Open. The Statement of Authorship is included. To end, I present the chapter summary and main findings identified from this study.

6.2 Dataset: Australian Rheumatology Association Database

The Australian Rheumatology Association Database (ARAD) is a voluntary national registry founded in 2001. The aims of the ARAD are to collect and to process prospective data on long-term safety and effectiveness of b/tsDMARDs and health outcomes in patients with inflammatory arthritis (RA, SpA, psoriatic arthritis, and juvenile idiopathic arthritis)[273, 274].

Study participation is on a voluntary basis and interested participants were recruited or referred by their treating rheumatologists across Australia. Information about this ARAD national registry was given to the participants. Eligible participants include those commencing on biologic DMARDs as well as those on csDMARDs (although the latter begun their enrolment since 2007). By agreeing to be part of this ARAD registry, there is no impact on the participants' treatment trajectory, as the participants may start, change, or stop their DMARDs at any timepoint during the ARAD study follow-up, as informed by their treating rheumatologists. In addition, there is no exclusion for disease duration, previous medication use, and comorbidities.

In this database, once consented, ARAD participants complete self-reported questionnaires biannually (options of paper or online format from August 2009) until January 2014. Thereafter, the questionnaires are completed once every year after the initial 2 years of biannual follow-up. The participant's disease diagnosis is determined based on expert opinion diagnosis (i.e., the treating rheumatologist) instead of classification criteria based.

Apart from diagnosis and seropositivity status (RF or ACPA), these self-reported questionnaires also consist of:

- Sociodemographic information – education, employment, marital status, private insurance, disability support
- Smoking history

- Alcohol consumption history
- Height and weight
- Comorbidities such as medical illnesses, infection, and malignancy
- Hospitalisations
- Current arthritis-related medication use – csDMARDs, b/tsDMARDs, glucocorticoids (oral/parenteral/intra-articular/intra-muscular), NSAIDs, analgesics, alternative medicine (herbal/complementary)
- Symptoms related to arthritis
- Measures of health outcomes and quality of life – HAQ-DI, SF-36, AQoL, European Quality of Life (EuroQoL)
- Global evaluation of disease activity VAS – pain and arthritis activity (in the past week)

Specific to the medication use, information on dose and duration of therapy is not collected in ARAD. Data on longitudinal safety outcomes in relation to DMARD use such as infection, malignancy and mortality is available through ARAD, and these outcomes can be used and validated with data linkage to other registries in Australia such as the Pharmaceutical Benefits Scheme (PBS), cancer registry, and death registry.

All ARAD participants provide written permission to be contacted by ARAD investigators and written informed consent for study participation and for anonymous data analyses and associated data linkages. Once enrolled and during the study follow-up, participants have the options to either opt out at any time, or alternatively, to be converted to 'tracking only' status, by which ongoing questionnaire is ceased but retrospective data collected can be used for research or data linkage as necessary. Ethics approval for ARAD has been granted by 18 committees and organisations across all Australian states and territories[273, 274].

6.2.1 Access to biologic/targeted synthetic DMARDs in Australia

In Australia, universal access to medications is provided under the PBS, which forms part of the Australian Government's broader National Medicines Policy[275]. Biologic DMARDs for inflammatory arthritis were only accessible through clinical trials prior to 2004. Thereafter, patients with inflammatory arthritis and with eligibility criteria fulfilled were able to access biologic DMARDs prescribed by their treating rheumatologists[276]. The eligibility criteria include patients with severe and active RA and with the following[276]:

- Failure to achieve disease control after at least 6 months of intensive csDMARDs (minimum of two agents used of at least 3 months for each, and includes methotrexate use, unless contraindicated) AND
- Serological markers showing active disease prior to biologic treatment – elevated ESR of >25mm/hour and/or elevated CRP of >15mg/L AND
- At least 20 active (swollen and tender) joints or at least 4 major active joints (elbow, wrist, knee, ankle, shoulder and/or hip)

For RA, currently approved b/tsDMARDs that are available through the PBS include abatacept (intravenous, IV or subcutaneous, SC), adalimumab, baricitinib, certolizumab, etanercept, golimumab, infliximab (IV or SC), tocilizumab (IV or SC), tofacitinib, upadacitinib, and rituximab (PBS approval is not required)[276].

6.3 Statement of Authorship

Statement of Authorship

Title of paper	Trajectories of Self-Reported Pain-Related Health Outcomes and Longitudinal Effects on Medication Use in Rheumatoid Arthritis: A Prospective Cohort Analysis Using the Australian Rheumatology Association Database (ARAD)
Publication status	Published
Publication Details	Pisaniello, H.L., Lester, S., Russell, O., Black, R.J., Tieu, J., Richards, B., Barrett, C., Lassere, M., March, L., Buchbinder, R., Whittle, S.L., Hill, C.L. Trajectories of Self-Reported Pain-Related Health Outcomes and Longitudinal Effects on Medication Use in Rheumatoid Arthritis: A Prospective Cohort Analysis Using the Australian Rheumatology Association Database (ARAD) RMD Open Impact Factor: 5.806

Principal Author

Name of principal author	Huai Leng Pisaniello
Contribution to the Paper	Contributed to the study design and conception, and data acquisition. Contributed to the data analysis and interpretation of study results. Contributed to the preparation, drafting and revision of the manuscript for publication.
Overall percentage (%)	75%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 8 April 2023

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. The candidate's stated contribution to the publication is accurate (as detailed above);
ii. Permission is granted for the candidate in include the publication in the thesis; and
iii. The sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Susan Lester
Contribution to the Paper	Contributed to the study design and conception, and data acquisition and management. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 18/04/2023

Name of Co-Author	Oscar Russell
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 15, 2023

Name of Co-Author	Rachel J. Black
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jul 4, 2023

Name of Co-Author	Joanna Tieu
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 15, 2023 Jun 15, 2023

Name of Co-Author	Bethan Richards
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 15, 2023

Name of Co-Author	Claire Barrett
Contribution to the Paper	Contributed to the interpretation of study results. Contributed ıe manuscript draft and approval of ɔlication.
Signature	Date: 7/8/23

Name of Co-Author	Marissa Lassere
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 10/11/2023

Name of Co-Author	Lyn March
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 14 August 2023

Name of Co-Author	Rachelle Buchbinder
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 15, 2023

Name of Co-Author	Samuel L. Whittle
Contribution to the Paper	Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	ate: 18/04/2023

Name of Co-Author	Catherine L. Hill
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.

Signature Date: 18/04/2023

6.4 Manuscript: Trajectories of Self-Reported Pain-Related Health Outcomes and Longitudinal Effects on Medication Use in Rheumatoid Arthritis: A Prospective Cohort Analysis Using the Australian Rheumatology Association Database (ARAD)

Published in *RMD Open Journal*

RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL RESEARCH

Trajectories of self-reported pain-related health outcomes and longitudinal effects on medication use in rheumatoid arthritis: a prospective cohort analysis using the Australian Rheumatology Association Database (ARAD)

Huai Leng Pisaniello ,[1,2] Susan Lester,[2,3] Oscar Russell,[1,2,3] Rachel Black,[2,3] Joanna Tieu,[2,3] Bethan Richards,[4,5,6] Claire Barrett,[7,8] Marissa Lassere,[9,10] Lyn March,[11,12] Rachelle Buchbinder ,[13] Samuel L Whittle,[2,3,13] Catherine L Hill[2,3]

To cite: Pisaniello HL, Lester S, Russell O, *et al*. Trajectories of self-reported pain-related health outcomes and longitudinal effects on medication use in rheumatoid arthritis: a prospective cohort analysis using the Australian Rheumatology Association Database (ARAD). *RMD Open* 2023;**9**:e002962. doi:10.1136/rmdopen-2022-002962

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/rmdopen-2022-002962).

Received 23 December 2022
Accepted 7 July 2023

For numbered affiliations see end of article.

Correspondence to
Dr Huai Leng Pisaniello; huaileng.pisaniello@adelaide.edu.au

ABSTRACT

Objective To determine distinct trajectories of self-reported pain-related health status in rheumatoid arthritis (RA), their relationship with sociodemographic factors and medication use.

Methods 988 Australian Rheumatology Association Database participants with RA (71% female, mean age 54 years, mean disease duration 2.3 years) were included. Distinct multi-trajectories over 15-year follow-up for five different self-reported pain-related health outcome measures (Health Assessment Questionnaire Disability Index, visual analogue scores for pain, arthritis, global health and the Assessment of Quality of Life utility index) were identified using latent variable discrete mixture modelling. Random effects models were used to determine associations with medication use and biologic therapy modification during follow-up.

Results Four, approximately equally sized, pain/health status groups were identified, ranging from 'better' to 'poorer', within which changes over time were relatively small. Important determinants of those with poorer pain/health status included female gender, obesity, smoking, socioeconomic indicators and comorbidities. While biologic therapy use was similar between groups during follow-up, biologic therapy modifications ($p_{linear}<0.001$) and greater tendency of non-tumour necrosis factor inhibitor use ($p_{linear}<0.001$) were observed in those with poorer pain/health status. Similarly, greater use of opioids, prednisolone and non-steroidal anti-inflammatory drugs was seen in those with poorer pain/health status.

Conclusion In the absence of disease activity information, distinct trajectories of varying pain/health status were seen from the outset and throughout the disease course in this RA cohort. More biologic therapy modifications and greater use in anti-inflammatories, opioids and prednisolone were seen in those with poorer pain/health status, reflecting undesirable lived experience of persistent pain in RA.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Despite treatment advances in rheumatoid arthritis (RA), persistent pain and suboptimal health status may remain significant in some patients, even in those with adequately controlled disease or in disease remission.

WHAT THIS STUDY ADDS

⇒ We performed multi-trajectory analysis, using five different self-reported pain-related health outcome measures, with patients with RA classified into four distinct pain-related health status subgroups, which were associated with sociodemographic and lifestyle factors.

⇒ Differences in the pain-related health status between subgroups were evident at baseline and were relatively stable over time. Greater use of opioids, anti-inflammatories and prednisolone and changes in biologic therapy were seen in those with poorer pain-related health status.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study highlights the importance of evaluating the overall well-being and pain experience of those living with RA from the outset, with a view to the development of appropriate management strategies in addition to suppression of disease inflammation.

INTRODUCTION

Over the last two decades, significant improvements in disease-related outcomes in rheumatoid arthritis (RA) are evident.[1–3] Patients presenting with inflammatory arthritis suspicious of RA are diagnosed earlier and treated intensively, alongside the major advances in targeted therapy using biologic/targeted synthetic disease modifying anti-rheumatic drugs (b/tsDMARDs).[1 2 4–6] However, mismatch between low disease activity or disease remission and patient-reported outcomes such as pain, fatigue and global disease activity remains. This is an ongoing treatment conundrum in the context of a treat-to-target approach in RA, especially in those with persistent pain despite remission in disease inflammation.[7 8]

Conventionally, the 28-Joint Disease Activity Score (DAS28), a universal composite scoring tool, is commonly used by treating physicians and in clinical trials to assess disease activity and treatment response in RA.[9 10] However, careful interpretation of the DAS28 scoring is crucial when it comes to determining disease remission objectively. For instance, discordance between the objective clinical assessment of joint inflammation and patient global disease activity (PGA) in the DAS28 scoring was observed in one-third of a multi-ethnic adult RA study cohort on treatment.[11] Patients with RA who have achieved a state of DAS28 remission may still experience clinically significant pain.[7] Similarly, McWilliams and colleagues identified 58% and 27% of their study cohort had partial improvement in pain and worsening pain after 12 months respectively, as assessed by the change of the DAS28-P index (defined as 'the proportion of DAS28 contributed by the patient-reported components') over time.[12] We recently showed that persistently high DAS28-P index scores predicted poor treatment response in an Australian early RA cohort, reflecting risks of underdiagnosed non-inflammatory pain and unnecessary escalation of RA treatment.[13] Such phenomenon is also observed in those with RA disease remission assessed using other index-based criteria endorsed by the American College of Rheumatology (ACR) and the European Alliance of Associations for Rheumatology (EULAR).[14] As highlighted by Studenic and his colleagues, two-thirds of their RA outpatient cohort achieved 'near-remission' based on fulfilling three of four of the Boolean-based criteria, with PGA being the primary limiting variable in defining disease remission.[15] Additionally, a cross-sectional study in Portugal demonstrated lack of ultrasonographic disease inflammation in RA patients with high levels of PGA reporting who were otherwise in remission.[16] More importantly, in a recent meta-analysis, despite using the ACR/EULAR Boolean-based remission criteria, patients in near-remission scored similarly in their PGA levels compared with those in non-remission, potentially with unintended risks of unnecessary treatment escalation.[17] These findings highlight the nuances behind the role of the PGA beyond the objective measure of disease remission in real-world clinical practice and the importance of dissecting the intention of treatment in RA, especially in those with persistent pain despite objective evidence of disease control.

Health status is regarded as the overall perception of the state of physical health, mental health and social well-being, and is an ever-changing metric in one's life course, ranging from a state of wellness to illness onset and its trajectory, if present.[18–20] In patients with RA, persistent pain may have an impact on their overall health status, as pain is regarded as the highest outcome priority for improvement.[2 21–23] Persistent pain in RA is multifaceted, largely driven not only by the complex dynamics between peripheral joint inflammation and nociplastic central pain processing, but it is also by the totality of the overall lived experience of the individual's health over time.[24 25] For example, higher levels of pain experienced in RA were significantly correlated with poor quality of life, defined by reduced overall health perception, lack of independence and decline in biopsychosocial functioning.[11 26–30] Therefore, looking into the intertwined relationship between the complex dynamics of pain in RA and the negative corollary health outcomes that followed has the potential to provide further insights into the overall impression of the well-being of the person living with RA. To date, little is known of the temporal relationship between pain and health status of patients living with RA, and more importantly, how pain-related health status trajectories translate into the patterns of medication use.

In this longitudinal study of a national cohort of patients with established RA, we first aimed to identify distinct subgroups of trajectories of self-reported pain-related health status, measured by different pain-related health outcome measures. Second, we aimed to examine the baseline sociodemographic and comorbidities within each of these identified trajectories. Finally, we aimed to use these identified trajectories as predictors of the time-varying effects on medication use.

METHODS

Study database

The Australian Rheumatology Association Database (ARAD) is a voluntary national registry founded in 2001 that aims to collect longitudinal data on long-term safety and effectiveness of b/tsDMARDs and health outcomes in patients with inflammatory arthritis. Information on the ARAD establishment, methodology and governance has been discussed in detail previously.[31 32]

In brief, ARAD participants with inflammatory arthritis completed self-reported questionnaires biannually (options of paper or online format from August 2009) until January 2014. Since then, questionnaires have been completed once every year after the initial 2 years of biannual follow-up. These self-reported questionnaires consist of sociodemographic information, current arthritis-related medication use, comorbidities (ie, comorbid medical illnesses), symptoms related to arthritis, and measures of health outcomes and quality of life.

All ARAD participants provided written permission to be contacted by ARAD investigators and written informed consent for study participation and for anonymous data analyses and associated data linkages. Ethics approval for ARAD has been granted by 18 committees snd organisations across all Australian states and territories.

Eligibility criteria

Study participants were selected from an ARAD snapshot from August 2021. Inclusion criteria were participants with rheumatologist-diagnosed RA and aged between 25 and 75 years old at diagnosis. We included ARAD participants who entered ARAD within 5 years of diagnosis and with at least 3 years of follow-up, therefore the study comprised participants with their RA diagnosis date between 1998 and 2018. Baseline data on age, gender, smoking and alcohol history, body mass index (BMI), education, employment, disability and comorbidities were extracted from the self-reported responses provided at ARAD entry.

Study outcomes

In this study, for included ARAD participants, five different self-reported pain-related health outcome measures derived from the ARAD questionnaires were used in the analysis, encompassing the overall pain experience and health status over time. These were (1) arthritis-related disability measured by the Health Assessment Questionnaire Disability Index, HAQ-DI (0–3 scale, higher score indicates higher level of disability), (2) pain level over the past week, measured on a 0–100 mm visual analogue score, VAS scale (higher score indicates greater pain), (3) participant-reported arthritis 'condition' (disease impact) measured on a 0–100 mm scale (higher score indicates worse arthritis), (4) participant-reported global health item measured on a 0–100 mm scale (higher score indicates better global health) and (5) the utility composite score of the Assessment of Quality of Life, AQoL, which ranges from 1.00 (indicating full health), to 0.00 (indicating death-equivalent), and to −0.04 (indicating a state worse than death).[33–36]

Socioeconomic status

In addition to education level and disability support reported in ARAD, socioeconomic status (SES) was also measured by the Index of Socioeconomic Advantage and Disadvantage (IRSAD), and Socio-Economic Indexes for Areas (SEIFA) developed by the Australian Bureau of Statistics.[37] The IRSAD quintile, according to 2016 Australian population census data, was assigned using SA1 areas, which are the smallest SEIFA, and correspond to an average of 400 people.

Comorbidity index

A modification of the Rheumatic Disease Comorbidity Index (RDCI) was used as a measure of comorbidity (range 0–9), with osteoporosis substituted for fracture.[38] Additional comorbidities in relation to medical illnesses that contribute to the index were lung disease, cardiovascular disease, hypertension, depression, cancer, gastrointestinal ulcer or stomach problems.

Medications

Current use of opioids, non-steroidal anti-inflammatory drugs (NSAIDs), oral glucocorticoids, oral and subcutaneous methotrexate, and b/tsDMARDs were obtained from each completed questionnaire, however, dosage information was not available.

Statistical analysis

All analyses were performed in Stata V.16.1 (StataCorp LLC, TS, USA).

Trajectories of self-reported pain-related health outcome measures

Group-based multi-trajectory modelling, using all five self-reported pain-related health outcome measures, was performed to identify distinct groups of participants followed from ARAD baseline to a maximum of 15 years. Analysis was performed using the Stata ado 'traj', which uses a discrete mixture modelling approach, a form of group-based trajectory modelling (GBTM), to stratify latent subgroups of participants based on homogeneity of between-individual trajectories for which follow-up time was modelled as a continuous variable, with both linear and quadratic terms.[39–41] The optimal number of trajectory groups was established based on the model selection criteria using the Akaike information criteria (AIC), Bayesian information criteria (BIC), entropy (which determines the overall probability of the individuals being accurately assigned to a homogenous trajectory) and the log-likelihood.[39–42] The reporting of this trajectory analysis was prepared in accordance with the Guidelines for Reporting on Latent Trajectory Studies Checklist.[43]

Baseline comparisons between the identified trajectory groups for sociodemographic, medication use and other relevant variables were performed using the Jonckheere-Terpstra test for ordinal data.

Medication use

Trajectory subgroups were considered as the predictors in a random intercept, longitudinal panel regression analysis of NSAIDs, glucocorticoids, opioids, methotrexate and b/tsDMARDs use over follow-up, using both binomial and multinomial (for b/tsDMARD use only) models. The results were interpreted as predicted marginal probabilities (with 95% CI) and orthogonal polynomial linear contrasts were used to assess ordinal trends between trajectory groups. The results were also presented as (subject-specific) ORs and the corresponding 95% CI, with all p values of <0.05 being considered statistically significant.

Modification of b/tsDMARD was also examined by time-to-event analysis in which failure times were defined at the initiation of, or change in, b/tsDMARD treatment. Multiple failures (modifications of b/tsDMARD) were possible for each individual, and therefore these data

Table 1 Baseline data on general demographics, socioeconomic demographics, medication use and comorbidities for participants stratified by pain-related health status trajectory groups

Baseline	All	Group 1 (better)	Group 2	Group 3	Group 4 (poorer)	p_{trend}
Number of participants, N	988	169	285	316	218	
Age at diagnosis: mean (SD)	53 (11)	50 (12)	52 (11)	54 (11)	54 (11)	0.005
Age at ARAD entry: mean (SD)	54 (11)	52 (12)	54 (11)	56 (11)	54 (11)	0.003
Disease duration (years): mean (SD)	2.3 (1.4)	2.3 (1.4)	2.2 (1.4)	2.4 (1.3)	2.4 (1.4)	0.17
Follow-up years: mean (SD)	6.7 (4.1)	6.9 (4.3)	7.2 (4.3)	6.7 (4.3)	5.9 (4.1)	0.004
Females: n (%)	703 (71%)	94 (56%)	200 (70%)	241 (76%)	168 (77%)	<0.001
BMI (WHO category): n (%)						<0.001
Normal	156/643 (24%)	40/121 (33%)	62/206 (30%)	40/194 (21%)	14/122 (11%)	
Overweight	228/643 (35%)	60/121 (50%)	74/206 (36%)	62/194 (32%)	32/122 (26%)	
Obese	259/643 (40%)	21/121 (17%)	70/206 (34%)	92/194 (47%)	76/122 (62%)	
Current smoker: n (%)	160/987 (16%)	18 (11%)	40 (14%)	61 (19%)	41 (19%)	0.009
Disability support: n (%)	111 (11%)	0	14 (5%)	33 (10%)	64 (29%)	<0.001
Education: n (%)						<0.001
Did not complete high school	245/987 (25%)	35 (21%)	62 (22%)	86 (27%)	62 (28%)	
Completed high school	337/987 (34%)	60 (36%)	77 (27%)	114 (36%)	86 (39%)	
Post high school	405/987 (41%)	74 (44%)	146 (51%)	115 (37%)	70 (32%)	
SES quintile**: n (%)						<0.001
Q1 (lowest)	163/831 (20%)	15/141 (11%)	40/229 (17%)	63/271 (23%)	45/90 (24%)	
Q2	173/831 (21%)	36/141 (25%)	38/229 (17%)	51/271 (19%)	48/90 (25%)	
Q3	178/831 (21%)	29/141 (21%)	41/229 (18%)	67/271 (25%)	41/90 (22%)	
Q4	160/831 (19%)	22/141 (16%)	56/229 (24%)	53/271 (20%)	29/90 (15%)	
Q5 (highest)	157/831 (19%)	39/141 (28%)	54/229 (24%)	37/271 (14%)	27/90 (14%)	
Comorbidity index: mean (SD)	1.0 (1.3)	0.6 (0.9)	0.8 (1.1)	1.1 (1.3)	1.8 (1.6)	<0.001
Trajectory analysis outcomes						
HAQ-DI: mean (SD)	1.0 (0.7)	0.4 (0.5)	0.7 (0.6)	1.2 (0.5)	1.7 (0.7)	<0.001
Pain VAS: mean (SD)	46 (26)	26 (24)	40 (23)	51 (23)	61 (21)	<0.001
Arthritis condition VAS: mean (SD)	46 (26)	27 (26)	40 (25)	53 (22)	61 (21)	<0.001
Global health VAS: mean (SD)	63 (20)	76 (19)	69 (17)	60 (17)	48 (19)	<0.001
AQoL utility index: mean (SD)	0.52 (0.25)	0.74 (0.21)	0.63 (0.18)	0.48 (0.19)	0.26 (0.17)	<0.001
Medications: n (%)						
Opioids	330 (33%)	25 (15%)	70 (25%)	121 (38%)	114 (52%)	<0.001
Prednisolone	471 (48%)	66 (39%)	141 (49%)	146 (46%)	118 (54%)	0.022
NSAIDs	444 (45%)	72 (43%)	138 (48%)	147 (47%)	87 (40%)	0.37
Methotrexate	723 (73%)	141 (83%)	219 (77%)	210 (66%)	153 (70%)	<0.001
Other csDMARD	20 (2%)	1 (0.6%)	4 (1.4%)	9 (29%)	6 (2%)	
b/tsDMARD	537 (54%)	91 (54%)	148 (52%)	167 (53%)	131 (60%)	0.17

Trend tests (p_{trend}) were performed using the Jonckheere-Terpstra test.
*SES was measured by the Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD).
AQoL, Assessment of Quality of Life; ARAD, Australian Rheumatology Association Database; BMI, body mass index; b/tsDMARDs, biologic/targeted synthetic disease modifying anti-rheumatic drugs; csDMARDs, conventional synthetic disease modifying anti-rheumatic drugs; HAQ-DI, Health Assessment Questionnaire Disability Index; NSAIDs, non-steroidal anti-inflammatory drugs; SES, socioeconomic status; VAS, visual analogue score.

Table 2 Model fit selection criteria for choosing the optimal number of trajectory latent classes (groups), and the number of participants assigned to each class

N_classes	AIC	BIC	LL	Entropy	Class 1 (%)	Class 2 (%)	Class 3 (%)	Class 4 (%)	Class 5 (%)
1	146 116	146 165	146 096		100	–	–	–	–
2	135 720	135 808	135 684	0.95	49	51	–	–	–
3	132 739	132 866	132 687	0.92	35	37	28	–	–
4	131 391	131 558	131 323	0.88	17	29	32	22	–
5	Inestimable (singular variance-covariance matrix)			–	–	–	–	–	–

Lower values of the Akaike information criterion (AIC), the Bayesian information criterion (BIC) and the log-likelihood (LL) indicate better model fit. Entropy is the average posterior probability of class membership, with values closer to one indicating greater precision, and values >0.7 indicating satisfactory discrimination between classes.

were analysed by a random effects, parametric Weibull 'survival' model, which models the baseline hazard rate and allows for within-individual dependencies between treatment failure episodes. The Weibull survival model had both proportional-hazards (PH) and accelerated failure-time (AFT) parameterisations, and both were reported. Regression coefficients for the PH model were expressed as HRs, with values of >1 indicating an increased risk of 'failure' occurring compared with the reference group at any given time point. For the AFT model, the regression coefficients were expressed as time ratios (TRs), with values of <1 indicating shorter b/tsDMARD 'failure' times.

RESULTS

A total of 988 ARAD participants were included in the study, the majority of whom were of Caucasian ancestry (93%) and spoke English at home (98%). Participants were predominantly female (71%) with a mean age at ARAD entry of 54 years (SD of 11) and a mean disease duration of 2.3 years (SD 1.4). The mean ARAD follow-up time was 6.7 years (maximum 15 years), as outlined in table 1.

Description of trajectories

Using multi-trajectory modelling, study participants were stratified into four, approximately equally sized, distinct pain-related health status groups. Four subgroups were selected based on the best model fit (minimum AIC, BIC and log-likelihood criteria) and a high entropy score (the average posterior probability of class membership) (table 2). Additional information on the model output was reported in online supplemental file.

The fitted multi-trajectories over time for each outcome measure for each group were reported in figure 1. The major difference between the groups was readily identifiable as the location (level) of the scores for each outcome, rather than the shape of the trajectories over time, and in fact, changes over time within each group were relatively small. Importantly, the patterns across each of the five outcome measures were remarkably similar, indicating that they measure the same underlying (latent) pain-related health construct. The four subgroups of study participants were therefore interpreted as an ordered classification of pain-related health status ranging from 'better pain-related health status' in group 1 to 'poorer pain-related health status' in group 4.

Baseline comparisons

Baseline comparisons between the four pain-related health status groups were reported in table 1. There was an increasing female predominance with poorer pain/health status (from 56% in group 1 to 77% in group 4, p<0.001), and a relatively small, but statistically significant, trend for baseline age (from 52 years in group 1 to 54 years in group 4, p=0.003). Importantly, the disease duration at ARAD entry was comparable across all four groups. In terms of other sociodemographic variables, obesity, current smoking, comorbidity index and lower SES indicators (such as education level, being on disability support and IRSAD quintile) were all associated with poorer pain/health status. Of note, 30% of those with poorer pain-related health status had self-reported diagnosis of depression, which was significantly higher than those in the better pain-related health status group (3%) (online supplemental table 1).

In detail, the five pain-related health outcome measures used for the multi-trajectory analysis were each different at baseline between the four groups of participants. When comparing the 'better' (group 1) to 'poorer' (group 4) groups at baseline, the HAQ-DI increased from 0.4 to 1.7, pain VAS from 26 to 61, arthritis condition VAS from 27 to 61, whereas the global health VAS decreased from 76 to 48 and the AQoL utility index decreased from 0.74 to 0.26 (all p<0.001). This was accompanied by an increase in baseline opioid use (from 15% to 52%), prednisolone use (from 39% to 54%), and perhaps surprisingly, lower methotrexate use (from 83% to 70%), which was possibly offset by a statistically non-significant increase in b/tsDMARD use (from 54% to 60%, p=0.17). However, differences in medication use between the pain-related health status groups was subsequently explored in detail over the duration of follow-up.

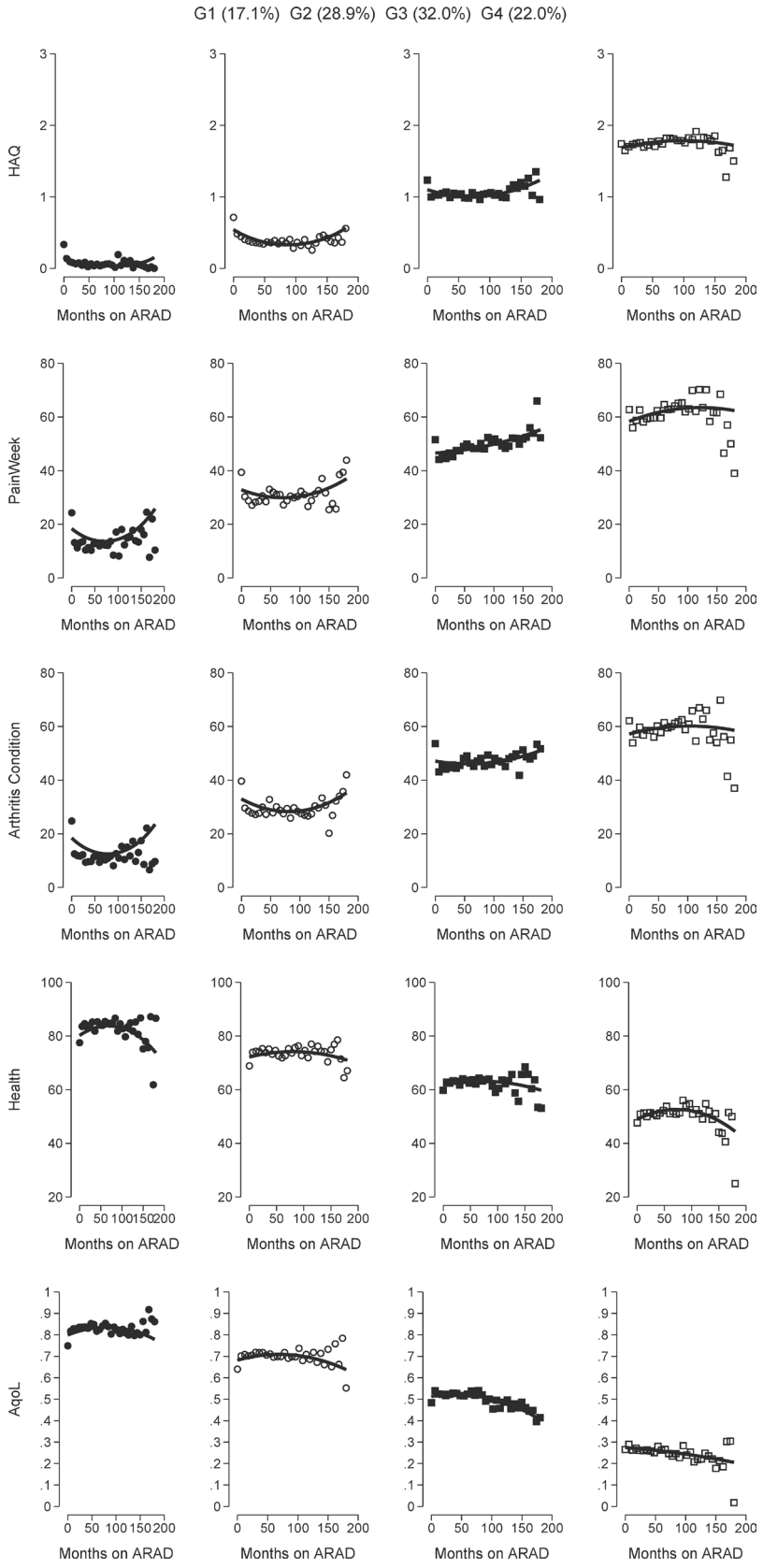


Figure 1 Changes in five self-reported pain-related health status outcomes in rheumatoid arthritis during ARAD follow-up. The outcomes (top to bottom panels) were the Health Assessment Questionnaire Disability Index (HAQ-DI), a pain visual analogue score (VAS), an arthritis condition VAS, patient's global health assessment and the Assessment of Quality of Life (AQoL) utility index. Four different longitudinal trajectory groups were identified (left to right panels). Each panel depicts the fitted regression line (estimated with both linear and quadratic terms for follow-up time) and mean estimates (symbols) at follow-up times for each outcome for each trajectory group. ARAD, Australian Rheumatology Association Database.

Table 3 Longitudinal random effect panel regression modelling analysis of the four pain-related health status trajectory groups as predictors for medication use during Australian Rheumatology Association Database follow-up

Trajectory group	Marginal probability (95% CI)	OR (95% CI)	P value
Opioid use: p_{linear}<0.001			
Group 1 (better)	0.08 (0.05 to 0.10)	1 (base)	
Group 2	0.18 (0.15 to 0.21)	4.7 (2.6 to 8.7)	<0.001
Group 3	0.36 (0.33 to 0.40)	24.9 (13.7 to 45.4)	<0.001
Group 4 (poorer)	0.57 (0.52 to 0.62)	116.1 (61.0 to 221.1)	<0.001
Prednisolone use: p_{linear}<0.001			
Group 1 (better)	0.26 (0.21 to 0.30)	1 (base)	
Group 2	0.36 (0.32 to 0.41)	3.3 (1.5 to 7.2)	0.003
Group 3	0.44 (0.39 to 0.48)	7.5 (3.4 to 16.2)	<0.001
Group 4 (poorer)	0.53 (0.48 to 0.57)	23.0 (9.9 to 53.7)	<0.001
NSAID use: p_{linear}=0.003			
Group 1 (better)	0.27 (0.23 to 0.32)	1 (base)	
Group 2	0.39 (0.35 to 0.43)	2.7 (1.5 to 4.7)	0.001
Group 3	0.41 (0.37 to 0.45)	3.2 (1.8 to 5.6)	<0.001
Group 4 (poorer)	0.37 (0.32 to 0.42)	2.4 (1.3 to 4.3)	0.005
Methotrexate use: p_{linear}=<0.001			
Group 1 (better)	0.78 (0.74 to 0.82)	1 (base)	
Group 2	0.75 (0.71 to 0.78)	0.60 (0.28 to 1.33)	0.21
Group 3	0.65 (0.61 to 0.70)	0.20 (0.09 to 0.44)	<0.001
Group 4 (poorer)	0.68 (0.63 to 0.74)	0.28 (0.12 to 0.65)	0.003
Other csDMARDs use: p_{linear} p=0.001			
Group 1 (better)	0.01 (0.00 to 0.02)	1 (base)	
Group 2	0.02 (0.01 to 0.03)	3.7 (0.6 to 24.7)	0.18
Group 3	0.04 (0.03 to 0.06)	21.9 (3.6 to 134.4)	0.001
Group 4 (poorer)	0.04 (0.03 to 0.05)	15.2 (2.3 to 100.5)	0.005
b/tsDMARD use			
1. No b/tsDMARDs: p_{linear}=0.34			
Group 1 (better)	0.35 (0.30 to 0.39)	1 (base)	
Group 2	0.32 (0.29 to 0.35)	1 (base)	
Group 3	0.32 (0.29 to 0.36)	1 (base)	
Group 4 (poorer)	0.31 (0.27 to 0.35)	1 (base)	
2. TNF inhibitors: p_{linear}=0.002			
Group 1 (better)	0.56 (0.51 to 0.61)	1 (base)	
Group 2	0.55 (0.52 to 0.59)	1.14 (0.6 to 2.2)	0.69
Group 3	0.50 (0.46 to 0.53)	0.85 (0.44 to 1.62)	0.61
Group 4 (poorer)	0.48 (0.44 to 0.53)	0.86 (0.42 to 1.75)	0.68
3. Other b/tsDMARDs: p_{linear}<0.001			
Group 1 (better)	0.09 (0.06 to 0.12)	1 (base)	
Group 2	0.13 (0.10 to 0.15)	2.19 (0.95 to 5.05)	0.065
Group 3	0.18 (0.15 to 0.20)	4.47 (1.98 to 10.09)	<0.001
Group 4 (poorer)	0.21 (0.17 to 0.24)	6.87 (2.87 to 16.46)	<0.001

Results are reported as both marginal probabilities (frequencies) and ORs with 95% CIs. Significance values are derived from both orthogonal polynomial linear contrasts of the marginal probabilities (p_{linear}) reflecting an overall ordinal trend, and Wald tests for individual ORs (p value).
b/tsDMARDs, biologic/targeted synthetic disease modifying anti-rheumatic drugs; csDMARDs, conventional synthetic disease modifying anti-rheumatic drugs; NSAID, non-steroidal anti-inflammatory drug; TNF, tumour necrosis factor.

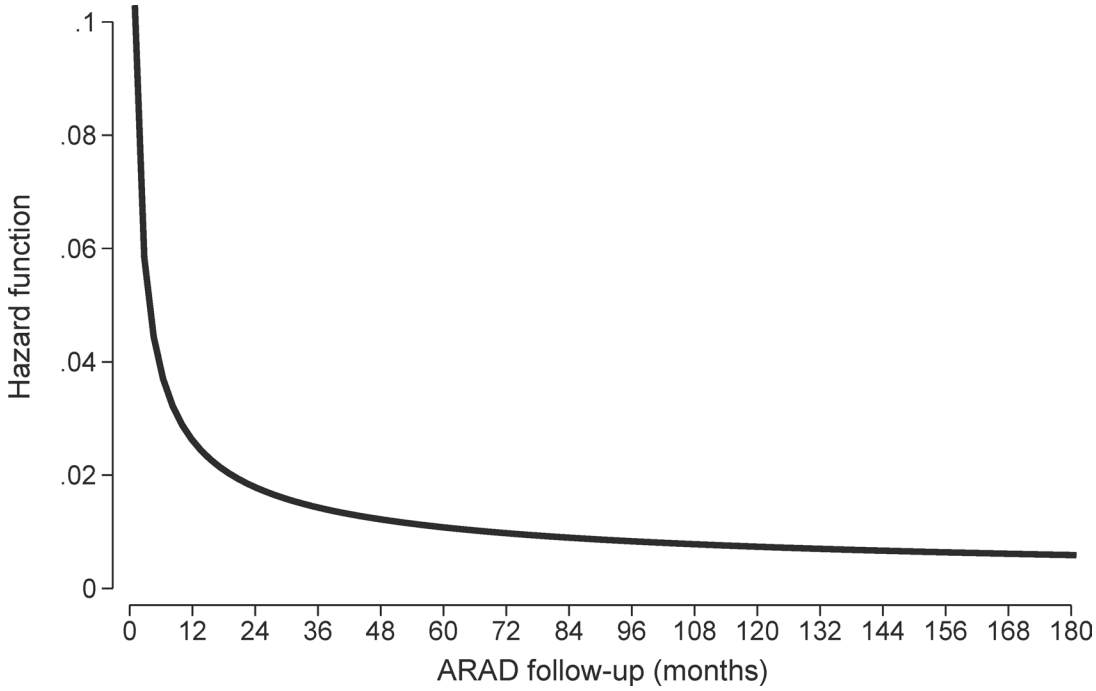


Figure 2 Marginal hazard rate for recurrent biologic/targeted synthetic disease modifying anti-rheumatic drug (b/tsDMARD) modification events estimated from a random effects Weibull parametric time-to-event proportional hazards model. The risk (hazard) of a b/tsDMARD treatment modification was greatest during the first 2 years or so following ARAD entry, and stabilised thereafter. ARAD, Australian Rheumatology Association Database.

Longitudinal (panel) mixed model regression analysis of medication use

Table 3 outlined the time-varying differences in medication use between the pain-related health status trajectory groups during ARAD follow-up.

Opioid, prednisolone and NSAIDs use each increased across the four pain-related health status groups. The difference in opioid use was the most marked, with the marginal probability increasing from 0.08 in the 'better pain-related health status' group (group 1) to 0.57 in the 'poorer pain-related health status' group (group 4) (p_{linear}<0.001). The marginal probability for prednisolone use was overall quite high, with an increase from 0.26 to 0.53 (p_{linear}<0.001).

In terms of DMARD use, the trend towards lower methotrexate use in those with poorer pain-related health status observed at baseline continued during follow-up with marginal probability ranging from 0.78 to 0.68 (p_{linear}<0.001). Although the use of other conventional synthetic DMARDs was low overall, the use of these medications increased with poorer pain-related health status. Overall, b/tsDMARD use was comparable across the four pain-related health status groups (p=0.34) but varied by type of b/tsDMARD. Specifically, tumour necrosis factor (TNF) inhibitor use decreased with poorer pain-related health status (marginal probability decreased from 0.56 to 0.48, p_{linear}=0.002), which was compensated by an increase in the use of other b/tsDMARDs (marginal probability increased from 0.09 to 0.21, p_{linear}<0.001).

Time-to-event analysis of b/tsDMARD modification

A total of 1567 b/tsDMARD modification episodes were identified for 988 participants in this analysis, with a median number of 2 episodes. The underlying hazard rate for b/tsDMARD modification (figure 2) indicated that the risk of b/tsDMARD treatment modification was

Table 4 Time-to-event analysis of recurrent biologic/targeted synthetic disease modifying anti-rheumatic drug treatment modifications

Trajectory group	HR (95% CI)	Time ratio (95% CI)	P value
Group 1 (better)	1 (base)	1 (base)	
Group 2	1.20 (1.01 to 1.42)	0.67 (0.46 to 0.98)	0.041
Group 3	1.58 (1.35 to 1.87)	0.36 (0.25 to 0.52)	<0.001
Group 4 (poorer)	1.78 (1.50 to 2.10)	0.28 (0.19 to 0.41)	<0.001

Analysis was performed by a random effects parametric Weibull time-to-event proportional hazards model, which may be parameterised as either an increased risk (HR) or accelerated failure time (time ratio).

highest within the first 2 years or so after ARAD study entry and plateaued thereafter.

Similar to our other results, there was an ordinal trend across the poorer pain-related health status groups, and the HR for b/tsDMARD modification for the 'poorer pain-related health status' (group 4) compared with the 'better pain-related health status' (group 1) was 1.78 (95% CI 1.50 to 2.10, p<0.001), as outlined in table 4. Alternatively, the AFT parameterisation indicated that the time to b/tsDMARD modification/failure was shorter by approximately 70% for group 4 participants compared with group 1 participants.

DISCUSSION

In this study of patients with rheumatologist-diagnosed RA, we used multi-trajectory analysis to identify subgroups of participants with an increasingly poorer pain-related health status. This type of analysis enabled us to examine risk factors, changes over time and medication use.

Our results highlight the strong interdependency between pain experience and overall health status in patients with RA. In part, pain experience in RA may be a proxy for the overall health status of the individuals. Notably, we observed these parallel patterns of synchronous trajectories of high pain and poor global health and more disability, and vice versa, from the outset and throughout the study follow-up period. Further, the changes over time within trajectory groups were minimal relative to the differences between groups, implying that DMARD treatment alone may not be sufficient to manage chronic pain in RA. However, there may be a window of opportunity early in the course of RA disease to identify patients at high risk of developing persistent pain.

Our study results showed that lifestyle factors, comorbidities and socioeconomic indicators, which are likely inter-related, were risk factors for a persistently poorer pain-related health status. These findings are similar to those from a large French observational study, which highlighted the temporal implications of pain heterogeneity and sociodemographic characteristics throughout the disease course.[44] Over the last two decades, the comorbidity burden at the time of RA diagnosis has risen, implying the need for early identification and better treatment tailoring for these at-risk individuals.[45] Additionally, our study findings have demonstrated high proportions of self-reported depression in those with poorer pain-related health status. The burden of pain in RA is highly correlated with levels of anxiety and depression, highlighting the unmet needs to consider open discussion of any psychological factors early on with these at-risk individuals, and to provide early psychosocial support or interventions as necessary.[30 46]

The use of opioids, prednisolone and NSAIDs throughout follow-up was higher in participants with worsening pain-related health status. The relationship with opioid use was the most marked, and consistent with a prior study of opioid use in the ARAD cohort which concluded that NSAID and DMARD treatment did not obviate opioid use in all patients.[47] Evidence for the benefits of opioid use in treating RA pain is minimal, resulting in a conditional recommendation against opioid use in the latest Australian Living Guideline for the treatment of inflammatory arthritis.[48] Concerningly, a recent American study has demonstrated that despite increasing awareness of the risks and harms associated with opioid use, chronic opioid use approximately doubled in patients with RA between 2002 and 2015, and was associated with pain, antidepressant use, high disease activity and disability.[49] In terms of prednisolone use, the Australian Living Guideline for the treatment of inflammatory arthritis recommends against long-term use of glucocorticoids in RA, and suggest aiming for the lowest dose and shortest possible duration of use of glucocorticoids when used for treatment of disease flare or as a bridging therapy when initiating DMARDs.[50] The relatively high probability of prednisolone use in this study, even in those with better pain-related health status, is potentially of concern. However, there is insufficient information in ARAD in terms of disease activity, prednisolone dose and duration to determine the appropriate prescribing of prednisolone in this study. Overarchingly, our study findings suggest that individuals with poorer pain-related health status did not experience abrogation of their pain level over time, despite greater use of opioids, prednisolone and anti-inflammatories. These at-risk individuals warrant further attentions when it comes to dissecting the underlying natural history of their pain experience, particularly in differentiating inflammatory and non-inflammatory pain in RA. In a proof-of-concept study by Wohlfahrt and her colleagues, lower knee pressure pain thresholds and conditioned pain modulation were shown to be predictive of DAS28 in those with low-moderate disease activity (pre-DMARD) and with higher baseline disease activity (post-DMARD), respectively.[51] Using these indices of pain sensitisation measures, in addition to the standard disease activity composite measures, may allow future personalised mechanism-specific pain interventions in RA, targeting those with PGA-near remission.[52]

The propensity of participants with RA and with poorer pain-related health status to have both used non-TNF inhibitor biological therapy, and experienced more b/tsDMARD treatment modifications, is consistent with more refractory and difficult-to-treat disease. Indeed, a prior study of ARAD participants identified that a lack of treatment response and side effects were the most common reasons for changing b/tsDMARDs, regardless of the line of treatment choice.[53] High pain level at the outset which persisted for up to 12 months was a strong predictor of discontinuation of TNF inhibitors, as shown in a British study of patients with RA.[54] Intriguingly, this was predominantly driven by the patient-reported pain/health components (as opposed to the inflammatory components) of the DAS28.[54]

In the current T2T strategy in managing patients with RA, the best approach to implement PGA in assessing

disease activity in RA remains controversial. Specifically, dilemma remains on how best to incorporate a comprehensive evaluation of the overall well-being and the patient-reported disease impact of individuals living with RA, distinct from the disease inflammation.[17] Although PGA is not necessarily a true reflection of biomarker of disease activity in RA, our study confirms the importance of early and consistent identification and intervention of pain-related health concerns in those at-risk individuals throughout the disease trajectory, as proposed in the current EULAR definition of 'difficult-to-treat' RA.[55] There is emerging evidence that the treatment response with regards to the self-reported pain/health (tender joint counts, global health) components of the DAS28 may be uncoupled from the inflammatory components response in some patients when assessing RA disease activity, suggesting a greater contribution of non-inflammatory factors, including central sensitisation, to pain in these patients.[12 13] In addition, a pragmatic dual-target strategy, focusing on disease inflammation and disease impact as separate composite indices, has been proposed to further refine the definition of disease remission in RA.[17 56 57] Focusing on capturing target information on disease impact, Patient Experienced Symptom State and seven items of Rheumatoid Arthritis Impact of Disease are some of the promising PGA tools that are feasible and universally acceptable for regular use in clinical practice.[17 58 59] Unfortunately, in our study, we were not able to evaluate this factor as disease activity information was not available in the ARAD cohort. Nonetheless, our study results indicate there is an unmet need to incorporate a careful well-being evaluation of patients with high pain and poor health status at diagnosis with the view to the development of appropriate management strategies in addition to suppression of inflammatory disease. Overarchingly, when implemented early from the outset of RA diagnosis, integrative health approaches such as psychological and social welfare access and support, interventions in physical activity and lifestyle factors, and management of comorbidities and related modifiable risk factors may influence the overall outlook of the health status of patients living with RA.[60 61] Timely use of these valuable integrative health strategies may promote more sustainable multidisciplinary care for patients with RA.

This was a long-term longitudinal study of pain-related health outcomes in a well-characterised, well-treated Australian RA cohort. Australia has universal healthcare, and all participants were under the care of a rheumatologist, with access to appropriate medications under the Pharmaceutical Benefits Scheme (PBS). Further, our cohort was homogenous in relation to ancestry and language. Therefore, confounding due to major inequities in access to, and navigation of healthcare were likely to be minimised. However, there are other limitations in our study. First, disease activity information was not available, limiting our study capability to track the relationship between disease activity and pain/health status. Second, medication use was self-reported. Although the accuracy of self-reported medication use by ARAD participants has been previously validated against data from the Australian PBS, dosage information and exact duration of medication use was not available.[62 63] Third, data on other non-inflammatory rheumatological diagnoses were not specifically captured in the ARAD dataset, and therefore, our study results may not be generalisable to patients with RA and other concomitant chronic pain conditions such as fibromyalgia. Fourth, in our trajectory analysis, although we did not perform any training, testing and validation of our study dataset, we have based our optimal model selection on the recommended standard model parameters required for trajectory study reporting, such as the use of AIC, BIC, log-likelihood criteria, entropy and the average posterior probability of class membership.[43]

In summary, poorer pain-related health status in patients with diagnosed RA in this ARAD cohort is associated with sociodemographic and lifestyle factors, and these time-varying factors do not appreciably improve during follow-up despite increased opioids, prednisolone and anti-inflammatory medication use as well as b/tsDMARD treatment modification. Early identification of those potentially at risk of worse prognosis in the context of persistently poorer pain-related health status in RA is necessary. Holistically, there is an unspoken requisite to consider the overall outlook of well-being in patients with RA when assessing disease activity and treatment response, ideally at the time of diagnosis and continuously throughout the disease course. Having better understanding of the evolution of health status in patients living with RA, alongside their pain experience, will fundamentally enrich the opportunities in providing high-quality patient-focused care.

Author affiliations
[1]Discipline of Medicine, The University of Adelaide, Adelaide, South Australia, Australia
[2]Rheumatology Research Group, Basil Hetzel Institute for Translational Health Research, The Queen Elizabeth Hospital, Woodville South, South Australia, Australia
[3]Deaprtment of Rheumatology, The Queen Elizabeth Hospital, Woodville South, South Australia, Australia
[4]Department of Rheumatology, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia
[5]Institute for Musculoskeletal Health, The University of Sydney and Sydney Local Health District, Sydney, New South Wales, Australia
[6]Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia
[7]Department of Medicine, The University of Queensland, Brisbane, Queensland, Australia
[8]Department of Rheumatology, Redcliffe Hospital, Redcliffe, Queensland, Australia
[9]Department of Medicine, University of New South Wales, Sydney, New South Wales, Australia
[10]Department of Rheumatology, St George Hospital, Kogarah, New South Wales, Australia
[11]Florance and Cope Professorial Department of Rheumatology, Royal North Shore Hospital, St Leonards, New South Wales, Australia
[12]Department of Rheumatology, Institute of Bone and Joint Research at Kolling Institute, University of Sydney, Sydney, New South Wales, Australia
[13]Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Victoria, Australia

Twitter Huai Leng Pisaniello @huaileng_jess, Susan Lester @Lester_Sue1, Oscar Russell @oscar_russell, Rachel Black @dr_rachelblack, Bethan Richards @BethanRichards3, Lyn March @lynmarch1, Rachelle Buchbinder @RachelleBuchbin, Samuel L Whittle @samwhittle and Catherine L Hill @CatherineL_Hill

Acknowledgements The authors sincerely acknowledge and thank the rheumatologists and ARAD participants for their ARAD data contribution. The authors would like to thank the ARAD Steering Committee and the ARAD Database staff members.

Contributors HLP, SEL, SLW and CLH contributed to the conception and design of the work, the acquisition, analysis and interpretation of the data. HLP contributed to the main drafting and the writing of the manuscript and all authors (HLP, SEL, OR, RB, JT, BR, CB, ML, LM, RB, SLW, CLH) contributed to the subsequent drafting and revision of the manuscript. All authors (HLP, SEL, OR, RB, JT, BR, CB, ML, LM, RB, SLW, CLH) read and approved the final manuscript. HLP is responsible for the overall content as the guarantor.

Funding The Australian Rheumatology Association Database (ARAD) receives support of unrestricted educational grants administered through the Australian Rheumatology Association from Pfizer Australia, AbbVie Pty Ltd, Eli Lilly Australia Pty Ltd, Sanofi Australia, Celgene Australian & NZ, Bristol Myers Squibb Australia Pty Ltd, Amgen Australia Pty Ltd, Aventis, AstraZeneca, and Roche. There is no current supporting funding. ARAD was previously supported by an Australian National Health and Medical Research Council (NHMRC) Enabling Grant (ID 384330). Infrastructure support for ARAD is provided by Cabrini Health, Monash University, Royal North Shore Hospital, and the Australian Rheumatology Association (ARA). This manuscript formed part of HLP's Ph.D. work, for which HLP was the recipient of the Ken Muirden Overseas Training Fellowship from the Arthritis Australia, an educational grant funded by the ARA, and HLP also received the Australia Postgraduate Scholarship from the University of Adelaide. RB is funded by an NHMRC Investigator Fellowship (APP1194483).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by ARAD has ethics approval obtained from Monash University and other sites, including Central Adelaide Local Health Network (CALHN). This study was granted approval by the ARAD Steering Committee. This study has ethics approval obtained from the CALHN Human Research Ethics Committee. ARAD participants provided informed consent to be enrolled in ARAD. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The dataset used and analysed specific for this study may be made upon reasonable request to the ARAD steering committee (https://arad.org.au).

ORCID iDs
Huai Leng Pisaniello http://orcid.org/0000-0002-0425-1697
Rachelle Buchbinder http://orcid.org/0000-0002-0597-0933

REFERENCES

1 Emery P. Early referral recommendation for newly diagnosed rheumatoid arthritis: evidence based development of a clinical guide. *Ann Rheum Dis* 2002;61:290–7.
2 Heiberg T, Finset A, Uhlig T, *et al*. Seven year changes in health status and priorities for improvement of health in patients with rheumatoid arthritis. *Ann Rheum Dis* 2005;64:191–5.
3 Sparks JA. Rheumatoid arthritis. *Ann Intern Med* 2019;170:ITC1–16.
4 Lacaille D, Avina-Zubieta JA, Sayre EC, *et al*. Improvement in 5-year mortality in incident rheumatoid arthritis compared with the general population-closing the mortality gap. *Ann Rheum Dis* 2017;76:1057–63.
5 Markusse IM, Akdemir G, Dirven L, *et al*. Long-term outcomes of patients with recent-onset rheumatoid arthritis after 10 years of tight controlled treatment: A randomized trial. *Ann Intern Med* 2016;164:523–31.
6 Uhlig T, Heiberg T, Mowinckel P, *et al*. Rheumatoid arthritis is milder in the new millennium: health status in patients with rheumatoid arthritis 1994-2004. *Ann Rheum Dis* 2008;67:1710–5.
7 Lee YC, Cui J, Lu B, *et al*. Pain persists in Das28 rheumatoid arthritis remission but not in ACR/EULAR remission: a longitudinal observational study. *Arthritis Res Ther* 2011;13:R83.
8 Nieuwenhuis WP, de Wit MP, Boonen A, *et al*. Changes in the clinical presentation of patients with rheumatoid arthritis from the early 1990s to the years 2010: earlier identification but more severe patient reported outcomes. *Ann Rheum Dis* 2016;75:2054–6.
9 Prevoo ML, van 't Hof MA, Kuper HH, *et al*. Modified disease activity scores that include twenty-eight-joint counts. development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44–8.
10 van der Heijde DM, van't Hof MA, van Riel PL, *et al*. Disease activity score. *Ann Rheum Dis* 1992;51:140.
11 Barton JL, Imboden J, Graf J, *et al*. Patient-physician Discordance in assessments of global disease severity in rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2010;62:857–64.
12 McWilliams DF, Zhang W, Mansell JS, *et al*. Predictors of change in bodily pain in early rheumatoid arthritis: an inception cohort study. *Arthritis Care Res* 2012;64:1505–13.
13 Pisaniello HL, Whittle SL, Lester S, *et al*. Using the derived 28-joint disease activity score patient-reported components (Das28-P) index as a discriminatory measure of response to disease-modifying anti-rheumatic drug therapy in early rheumatoid arthritis. *BMC Rheumatol* 2022;6:67.
14 Studenic P, Aletaha D, de Wit M, *et al*. American college of rheumatology/EULAR remission criteria for rheumatoid arthritis: 2022 revision. *Ann Rheum Dis* 2023;82:74–80.
15 Studenic P, Smolen JS, Aletaha D. Near misses of ACR/EULAR criteria for remission: effects of patient global assessment in Boolean and index-based definitions. *Ann Rheum Dis* 2012;71:1702–5.
16 Brites L, Rovisco J, Costa F, *et al*. High patient global assessment scores in patients with rheumatoid arthritis otherwise in remission do not reflect Subclinical inflammation. *Joint Bone Spine* 2021;88:105242.
17 Duarte C, Ferreira RJO, Santos EJF, *et al*. Treating-to-target in rheumatology: theory and practice. *Best Pract Res Clin Rheumatol* 2022;36:101735.
18 Callahan D. The WHO definition of 'health *Stud Hastings Cent* 1973;1:77–88.
19 Health Do. What is health? 2022. Available: https://www.aihw.gov.au/reports/australias-health/what-is-health
20 Rumsfeld JS. Health status and clinical practice: when will they meet *Circulation* 2002;106:5–7.
21 Heiberg T, Kvien TK. Preferences for improved health examined in 1,024 patients with rheumatoid arthritis: pain has highest priority. *Arthritis Rheum* 2002;47:391–7.
22 Minnock P, FitzGerald O, Bresnihan B. Women with established rheumatoid arthritis perceive pain as the predominant impairment of health status. *Rheumatology (Oxford)* 2003;42:995–1000.
23 Sanderson T, Morris M, Calnan M, *et al*. Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. *Arthritis Care Res (Hoboken)* 2010;62:647–56.
24 Vladimirova N, Jespersen A, Bartels EM, *et al*. Pain Sensitisation in women with active rheumatoid arthritis: A comparative cross-sectional study. *Arthritis* 2015;2015:434109.
25 Zhang A, Lee YC. Mechanisms for joint pain in rheumatoid arthritis (RA): from Cytokines to central sensitization. *Curr Osteoporos Rep* 2018;16:603–10.
26 Courvoisier DS, Agoritsas T, Glauser J, *et al*. Pain as an important Predictor of Psychosocial health in patients with rheumatoid arthritis. *Arthritis Care Res* 2012;64:190–6.
27 Katchamart W, Narongroeknawin P, Chanapai W, *et al*. Health-related quality of life in patients with rheumatoid arthritis. *BMC Rheumatol* 2019;3:34.
28 Malm K, Bergman S, Andersson ML, *et al*. Quality of life in patients with established rheumatoid arthritis: A Phenomenographic study. *SAGE Open Med* 2017;5:2050312117713647.

29 Martinec R, Pinjatela R, Balen D. Quality of life in patients with rheumatoid arthritis - a preliminary study. *Acta Clin Croat* 2019;58:157–66.
30 Vergne-Salle P, Pouplin S, Trouvin AP, *et al*. The burden of pain in rheumatoid arthritis: impact of disease activity and psychological factors. *Eur J Pain* 2020;24:1979–89.
31 Buchbinder R, March L, Lassere M, *et al*. Effect of treatment with biological agents for arthritis in Australia: the Australian rheumatology Association database. *Intern Med J* 2007;37:591–600.
32 Williams MP, Buchbinder R, March L, *et al*. The Australian rheumatology Association database (ARAD). *Semin Arthritis Rheum* 2011;40:e2–3.
33 Fries JF, Spitz P, Kraines RG, *et al*. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137–45.
34 Hawthorne G, Osborne RH, Taylor A, *et al*. The Sf36 version 2: critical analyses of population weights, scoring Algorithms and population norms. *Qual Life Res* 2007;16:661–73.
35 Hawthorne G, Richardson J, Osborne R. The assessment of quality of life (Aqol) instrument: a Psychometric measure of health-related quality of life. *Qual Life Res* 1999;8:209–24.
36 Rabin R, de Charro F. EQ-5D: a measure of health status from the Euroqol group. *Ann Med* 2001;33:337–43.
37 Statistics ABo. Census of population and housing: socio-economic indexes for areas (SEIFA). statistical area level 1, indexes 2016. 2016. Available: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/2033.0.55.0012016?OpenDocument
38 England BR, Sayles H, Mikuls TR, *et al*. Validation of the rheumatic disease Comorbidity index. *Arthritis Care & Research* 2015;67:865–72.
39 Jones BL, Nagin DS, Roeder K. A SAS procedure based on mixture models for estimating developmental Trajectories. *Sociological Methods & Research* 2001;29:374–93.
40 Jones BL, Nagin DS. Advances in group-based trajectory modeling and an SAS procedure for estimating them. *Sociological Methods & Research* 2007;35:542–71.
41 Jones BL. A note on a STATA Plugin for estimating group-based trajectory models. *Sociological Methods & Research* 2013;42:608–13.
42 Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. *Annu Rev Clin Psychol* 2010;6:109–38.
43 van de Schoot R, Sijbrandij M, Winter SD, *et al*. The Grolts-checklist: guidelines for reporting on latent trajectory studies. *Structural Equation Modeling: A Multidisciplinary Journal* 2017;24:451–67.
44 Kumaradev S, Roux C, Sellam J, *et al*. Socio-demographic determinants in the evolution of pain in inflammatory rheumatic diseases: results from ESPOIR and DESIR cohorts. *Rheumatology (Oxford)* 2022;61:1496–509.
45 Nikiphorou E, Norton S, Carpenter L, *et al*. Secular changes in clinical features at presentation of rheumatoid arthritis: increase in Comorbidity but improved inflammatory States. *Arthritis Care Res (Hoboken)* 2017;69:21–7.
46 Backman CL. Psychosocial aspects in the management of arthritis pain. *Arthritis Res Ther* 2006;8:221.
47 Black RJ, Richards B, Lester S, *et al*. Factors associated with commencing and ceasing opioid therapy in patients with rheumatoid arthritis. *Semin Arthritis Rheum* 2019;49:351–7.
48 An Australian living guideline for the pharmacological management of inflammatory arthritis: opioids for pain management. 2022. Available: https://app.magicapp.org/#/guideline/LqRV3n
49 Lee YC, Kremer J, Guan H, *et al*. Chronic opioid use in rheumatoid arthritis: prevalence and predictors. *Arthritis Rheumatol* 2019;71:670–7.
50 An Australian living guideline for the pharmacological management of inflammatory arthritis: glucocorticoid therapy as an adjunct to DMARD therapy in the treatment of rheumatoid arthritis. 2022. Available: https://app.magicapp.org/#/guideline/LqRV3n
51 Wohlfahrt A, Muhammad LN, Song J, *et al*. Pain mechanisms associated with disease activity in patients with rheumatoid arthritis treated with disease-modifying Antirheumatic drugs: A regression tree analysis. *J Rheumatol* 2023;50:741–7.
52 McWilliams DF, Walsh DA. Inflammatory and Noninflammatory disease activity in rheumatoid arthritis: the effect of pain on personalized medicine. *J Rheumatol* 2023;50:721–3.
53 Fletcher A, Lassere M, March L, *et al*. Patterns of biologic and targeted-synthetic disease-modifying Antirheumatic drug use in rheumatoid arthritis in Australia. *Rheumatology (Oxford)* 2022;61:3939–51.
54 McWilliams DF, Walsh DA. Factors predicting pain and early discontinuation of tumour necrosis factor-alpha-inhibitors in people with rheumatoid arthritis: results from the British society for rheumatology Biologics register. *BMC Musculoskelet Disord* 2016;17:337.
55 Nagy G, Roodenrijs NM, Welsing PM, *et al*. EULAR definition of difficult-to-treat rheumatoid arthritis. *Ann Rheum Dis* 2021;80:31–5.
56 Ferreira RJO, Duarte C, Ndosi M, *et al*. Suppressing inflammation in rheumatoid arthritis: does patient global assessment blur the target? A practice-based call for a paradigm change. *Arthritis Care Res (Hoboken)* 2018;70:369–78.
57 Ferreira RJO, Ndosi M, de Wit M, *et al*. Dual target strategy: a proposal to mitigate the risk of Overtreatment and enhance patient satisfaction in rheumatoid arthritis. *Ann Rheum Dis* 2019;78:e109.
58 Duarte C, Santos E, da Silva JAP, *et al*. The patient experienced symptom state (PESS): a patient-reported global outcome measure that may better reflect disease remission status. *Rheumatology (Oxford)* 2020;59:3458–67.
59 Duarte C, Santos EJF, Ferreira RJO, *et al*. Validity and reliability of the EULAR instrument RAID.7 as a tool to assess individual domains of impact of disease in rheumatoid arthritis: a cross-sectional study of 671 patients. *RMD Open* 2021;7:e001539.
60 Gwinnutt JM, Wieczorek M, Balanescu A, *et al*. EULAR recommendations regarding lifestyle Behaviours and work participation to prevent progression of rheumatic and musculoskeletal diseases. *Ann Rheum Dis* 2023;82:48–56.
61 Rausch Osthoff A-K, Niedermann K, Braun J, *et al*. EULAR recommendations for physical activity in people with inflammatory arthritis and osteoarthritis. *Ann Rheum Dis* 2018;77:1251–60.
62 Lynch T, Buchbinder R, *et al*. Predictors of discordant prescription medication self-report in the Australian rheumatology Association database (ARAD) cohort. Australian Rheumatology Association 62nd Annual Scientific Meeting 20-22 May 2022; 2022
63 Lynch T, Barrett C. Assessing self-reported prescription medication validity in the Australian rheumatology Association database (ARAD) using pharmaceutical benefits scheme (PBS) claims data ARA-NSW 2021. 43rd Annual NSW Branch Meeting; Sydney, NSW, Australia, 2021

Supplementary File

[Trajectories of Self-Reported Pain-Related Health Outcomes and Longitudinal Effects on Medication Use in Rheumatoid Arthritis: A Prospective Cohort Analysis Using the Australian Rheumatology Association Database (ARAD)]

Four groups, 5 outcome trajectory model

988 observations read.
3 had no trajectory data in one or more models.
985 observations used in the trajectory model.

1.HAQ

Maximum Likelihood Estimates
Model: Censored Normal (cnorm)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	-0.06313	0.03255	-1.939	0.0525
	Linear	-0.01045	0.00122	-8.561	0.0000
	Quadratic	0.00006	0.00001	6.610	0.0000
2	Intercept	0.49903	0.02125	23.489	0.0000
	Linear	-0.00630	0.00073	-8.689	0.0000
	Quadratic	0.00004	0.00000	7.220	0.0000
3	Intercept	1.09758	0.01960	55.994	0.0000
	Linear	-0.00274	0.00069	-3.964	0.0001
	Quadratic	0.00002	0.00000	4.043	0.0001
4	Intercept	1.68822	0.02343	72.060	0.0000
	Linear	0.00200	0.00090	2.215	0.0268
	Quadratic	-0.00001	0.00001	-1.519	0.1288
	Sigma	0.50134	0.00447	112.256	0.0000

2. PainWeek

Maximum Likelihood Estimates
Model: Censored Normal (cnorm)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	15.60993	1.07749	14.487	0.0000
	Linear	-0.18391	0.03921	-4.691	0.0000
	Quadratic	0.00130	0.00027	4.787	0.0000
2	Intercept	32.32869	0.87359	37.007	0.0000
	Linear	-0.09043	0.02954	-3.061	0.0022
	Quadratic	0.00064	0.00020	3.138	0.0017
3	Intercept	46.52522	0.78020	59.632	0.0000
	Linear	0.01334	0.02808	0.475	0.6348
	Quadratic	0.00020	0.00020	1.013	0.3112
4	Intercept	58.55165	0.96685	60.559	0.0000
	Linear	0.08709	0.03723	2.339	0.0193
	Quadratic	-0.00036	0.00027	-1.296	0.1952
	Sigma	20.83857	0.15436	134.998	0.0000

3.Arthcond

Maximum Likelihood Estimates
Model: Censored Normal (cnorm)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	15.55170	1.10300	14.099	0.0000
	Linear	-0.21834	0.04003	-5.455	0.0000
	Quadratic	0.00141	0.00028	5.087	0.0000
2	Intercept	32.42543	0.89413	36.265	0.0000
	Linear	-0.12636	0.03016	-4.190	0.0000
	Quadratic	0.00078	0.00021	3.758	0.0002
3	Intercept	47.05029	0.79555	59.142	0.0000
	Linear	-0.03811	0.02879	-1.324	0.1855
	Quadratic	0.00034	0.00020	1.680	0.0929
4	Intercept	57.36250	0.98789	58.066	0.0000
	Linear	0.05893	0.03782	1.558	0.1192
	Quadratic	-0.00028	0.00028	-1.023	0.3064
	Sigma	21.25620	0.15766	134.821	0.0000

4. Health

Maximum Likelihood Estimates
Model: Censored Normal (cnorm)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	81.29972	0.85919	94.623	0.0000
	Linear	0.12357	0.03110	3.973	0.0001
	Quadratic	-0.00091	0.00021	-4.224	0.0000
2	Intercept	72.47985	0.66679	108.699	0.0000
	Linear	0.05356	0.02322	2.307	0.0211
	Quadratic	-0.00034	0.00016	-2.113	0.0346
3	Intercept	61.80468	0.62737	98.514	0.0000
	Linear	0.03977	0.02275	1.748	0.0804
	Quadratic	-0.00028	0.00016	-1.791	0.0734
4	Intercept	48.91202	0.76541	63.903	0.0000
	Linear	0.10332	0.02958	3.493	0.0005
	Quadratic	-0.00071	0.00022	-3.280	0.0010
	Sigma	16.45292	0.12034	136.720	0.0000

5.Aqol

Maximum Likelihood Estimates
Model: Censored Normal (cnorm)

Group	Parameter	Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
1	Intercept	0.80817	0.00856	94.448	0.0000
	Linear	0.00103	0.00030	3.413	0.0006
	Quadratic	-0.00001	0.00000	-3.083	0.0021
2	Intercept	0.68344	0.00643	106.267	0.0000
	Linear	0.00081	0.00022	3.650	0.0003
	Quadratic	-0.00001	0.00000	-3.875	0.0001
3	Intercept	0.51275	0.00619	82.881	0.0000
	Linear	0.00053	0.00022	2.468	0.0136

	Quadratic	-0.00001	0.00000	-4.187	0.0000
4	Intercept	0.27277	0.00746	36.575	0.0000
	Linear	-0.00020	0.00028	-0.711	0.4770
	Quadratic	-0.00000	0.00000	-0.523	0.6011
	Sigma	0.15819	0.00120	131.528	0.0000

Group membership

1	(%)	17.13935	1.27724	13.419	0.0000
2	(%)	28.85902	1.53871	18.755	0.0000
3	(%)	31.97457	1.56734	20.401	0.0000
4	(%)	22.02707	1.35673	16.235	0.0000

BIC=-131690.88 (N=49352) BIC=-131557.80 (N=985) AIC=-131391.45 ll= -131323.45

Entropy = 0.881

Supplementary Table 1: Baseline data on comorbidities (comorbid medical illnesses) for ARAD participants stratified by trajectories of pain-related health status

Baseline	Group1 (better)	Group2	Group3	Group4 (poorer)	P-value	Total
Number of participants, *N*	169	285	316	218		988
COMORBIDITIES: n (%)						
Hypertension	28 (17%)	58 (20%)	77 (24%)	70 (32%)	<0.001	233 (24%)
Hypercholesterolemia	15 (9%)	37 (13%)	55 (17%)	30 (14%)	0.085	137 (14%)
Cardiovascular diseases[1]	0	4 (1.4%)	5 (1.6%)	12 (5.5%)	< 0.001	21 (2%)
Diabetes mellitus[2]	9 (5%)	21 (7%)	27 (9%)	24 (11%)	0.038	81 (8%)
Thyroid disease	7 (4%)	24 (8%)	33 (10%)	28 (13%)	0.003	92 (9%)
Lung diseases[3]	9 (5%)	21 (7%)	36 (11%)	41 (19%)	<0.001	107 (11%)
Gastrointestinal diseases[4]	9 (5%)	36 (13%)	56 (18%)	49 (22%)	<0.001	150 (15%)
Osteoporosis	5 (3%)	23 (8%)	32 (10%)	23 (11%)	0.008	83 (8%)
Depression	5 (3%)	28 (10%)	41 (13%)	66 (30%)	< 0.001	140 (14%)

ARAD: Australian Rheumatology Association Database,
[1]Cardiovascular diseases include angina (stable/unstable), heart attack, coronary artery bypass graft and coronary angioplasty/stenting.
[2]Diabetes mellitus include insulin- and non-insulin dependent types.
[3]Lung diseases include asthma, bronchiectasis, and smoking-related disease.
[4]Gastrointestinal diseases include ulcers, reflux, hiatus hernia, Crohn's disease, and ulcerative colitis.

6.5 Summary

Pain-related health outcome measures are important determinants that reflect the overall well-being and pain experience in patients with RA. Using the 15-year follow-up ARAD data of five different self-reported pain-related health outcome measures (Health Assessment Questionnaire Disability Index, visual analogue scores for pain, arthritis, global health, and the Assessment of Quality of Life utility index), four trajectories of varying pain-related health status, ranging from 'better' to 'poorer', were identified. In this trajectory analysis, pain experience and health status were found to be strongly interdependent in RA.

In those with poorer pain-related health status, several main findings were observed:

- Significant sociodemographic indicators – female predominance, obesity, current smoking, more comorbidities (including depression) and lower socioeconomic status (including lower education level and disability support use)
- Significantly greater use of opioid, prednisolone and NSAID use
- More biologic modifications (despite a homogenous pattern of biologic use between groups during follow-up) and greater tendency of non-tumour necrosis factor inhibitor use

Early identification of patients with high pain and poor health status at the time of diagnosis of RA is necessary. This initial step may provide a window of opportunity to intervene early with psychosocial support and management of comorbidities and other modifiable risk factors.

To better understand the characteristics of individuals with persistent use of opioid, important determinants such as opioid-induced hyperalgesia or opioid-related sensitisation should be considered, which are beyond the information available in the ARAD dataset. Additionally, pain-related psychosocial determinants (such as catastrophising, coping, self-efficacy, depression, anxiety, and stress) of those with poorer pain-related health status are

important considerations in the study analysis, of which such data is not completely available in the ARAD and therefore, limiting further sub analysis of these psychosocial factors.

Chapter 7: Trajectories of Pain-Related Health Status in RA and the Associations with Hospitalisations and Mortality Risk

7.1 Preface

The study presented in this chapter is an extension from the research work done in Chapter 6. In this study, using the identified trajectory subgroups of pain-related health outcomes as predictors, as captured by the same clustering method used in study presented in Chapter 6, the time-varying effects on hospitalisations, mortality risk and causes of death were examined.

This study is written as a short clinical paper and is presented in a manuscript format in this chapter, and it has been submitted as a Letter to a peer-reviewed journal, Annals of the Rheumatic Diseases. As it is submitted as a Letter, I have written this study with a word limit of 600 and a reference limit of 6 in accordance with the journal submission guideline. The numbered references in this manuscript have been reformatted accordingly in this thesis, corresponding to the Reference thesis chapter. The Statement of Authorship is included. To end, I present the chapter summary and main findings identified from this study.

7.2 Statement of Authorship

Statement of Authorship

Title of paper	The Associations Between Poorer Pain-Related Health Status and Increased Hospitalisations and Excess Mortality in Patients with Rheumatoid Arthritis: A Prospective Cohort Analysis Using the Australian Rheumatology Association Database (ARAD)
Publication status	Submitted for publication
Publication Details	Pisaniello, H.L., Lester, S., Russell, O., Black, R.J., Tieu, J., Richards, B., Barrett, C., Lassere, M., March, L., Buchbinder, R., Hill, C.L., Whittle, S.L. The Associations Between Poorer Pain-Related Health Status and Increased Hospitalisations and Excess Mortality in Patients with Rheumatoid Arthritis: A Prospective Cohort Analysis Using the Australian Rheumatology Association Database (ARAD) Annals of Rheumatic Diseases Impact Factor: 28.5

Principal Author

Name of principal author	Huai Leng Pisaniello
Contribution to the Paper	Contributed to the study design and conception, and data acquisition. Contributed to the data analysis and interpretation of study results. Contributed to the preparation, drafting and revision of the manuscript for publication.
Overall percentage (%)	75%
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 8 April 2023

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. The candidate's stated contribution to the publication is accurate (as detailed above);
ii. Permission is granted for the candidate in include the publication in the thesis; and
iii. The sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Susan Lester
Contribution to the Paper	Contributed to the study design and conception, and data acquisition and management. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 18/04/2023

Name of Co-Author	Oscar Russell
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 15, 2023

Name of Co-Author	Rachel J. Black
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jul 4, 2023

Name of Co-Author	Joanna Tieu
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 15, 2023 Jun 15, 2023

Name of Co-Author	Bethan Richards
Contribution to the Paper	Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.

Date: Jun 15, 2023

Name of Co-Author	Claire Barrett
Contribution to the Paper	Contributed to the interpretation of study results. Contributed [illegible] he manuscript draft and approval of [illegible]plication.
Signature	Date: 7/8/23

Name of Co-Author	Marissa Lassere
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 10/11/2023

Name of Co-Author	Lyn March
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 14 August 2023

Name of Co-Author	Rachelle Buchbinder
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: Jun 15, 2023

Name of Co-Author	Catherine L. Hill
Contribution to the Paper	Contributed to the study data acquisition and management. Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.

Signature: Date: 18/04/2023

Name of Co-Author	Samuel L. Whittle
Contribution to the Paper	Contributed to the data analysis and interpretation of study results. Contributed to the critical appraisal of the manuscript draft and approval of the final manuscript for publication.
Signature	Date: 18/04/2023

For submission to the *Annals of the Rheumatic Diseases*

TITLE

The Associations Between Poorer Pain-Related Health Status and Increased Hospitalisations and Excess Mortality in Patients with Rheumatoid Arthritis: A Prospective Cohort Analysis Using the Australian Rheumatology Association Database (ARAD)

AUTHORS

Huai Leng Pisaniello MBBS, FRACP[1,2],

Susan Lester BSc[2,3],

Oscar Russell MBBS[1,2,3],

Rachel J. Black MBBS, PhD, FRACP[2,3],

Joanna Tieu MBBS, PhD, FRACP[2,3],

Bethan Richards MBBS (Hons), MMed (ClinEpi), PhD, FRACP[4,5,6],

Claire Barrett BSc (Hons), MBBS, FRCP, FRACP, GAICD[7,8],

Marissa Lassere MD, PhD[9,10],

Lyn March MBBS (Hons), MSc (Epi), PhD, FRACP, FAFPHM[11,12],

Rachelle Buchbinder MBBS (Hons), MSc, PhD, FRACP, FAHMS[13],

Catherine L. Hill MBBS, MD, MSc, FRACP[2,3],

Samuel L. Whittle MBBS (Hons), MClinEpi, FRACP[2,3,13]

AFFILIATIONS

[1]Discipline of Medicine, Faculty of Health and Medical Sciences, The University of Adelaide, South Australia, Australia.

[2]Rheumatology Research Group, Basil Hetzel Institute for Translational Health Research, The Queen Elizabeth Hospital, Woodville South, South Australia, Australia.

[3]Rheumatology Unit, The Queen Elizabeth Hospital, Woodville South, South Australia, Australia.

[4]Institute for Musculoskeletal Health, The University of Sydney and Sydney Local Health District, Sydney, New South Wales, Australia.

[5]Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia.

[6]Department of Rheumatology, Royal Prince Alfred Hospital, Sydney Local Health District, Sydney, New South Wales, Australia.

[7]Department of Rheumatology, Redcliffe Hospital, Redcliffe, Queensland, Australia.

[8]Department of Medicine, School of Medicine, University of Queensland, Brisbane, Queensland, Australia.

[9]Department of Rheumatology, St George Hospital, Kogarah, New South Wales, Australia.

[10]Department of Medicine, School of Population Health, University of New South Wales, New South Wales, Australia.

[11]Florance and Cope Professorial Department of Rheumatology, Royal North Shore Hospital, St Leonards, New South Wales, Australia.

[12]Department of Rheumatology, Institute of Bone and Joint Research at Kolling Institute, University of Sydney, Sydney, New South Wales, Australia.

[13]School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia.

ABSTRACT

Aim: Distinct trajectory groups of self-reported pain-related health states in patients with RA were identified, and their relationships with hospitalisation, mortality rates, and causes of death (COD) were examined.

Methods: 806 ARAD participants with RA and data linkage consent to the Australian Death Registry (72% female, mean age 54 years) were included. Distinct multi-trajectories for five different self-reported pain/health outcome measures (Health Assessment Questionnaire-Disability Index, visual analogue scores for pain, arthritis, global health, and the Assessment of Quality of Life-utility index) were identified using group-based trajectory modelling. Hospitalisation events were analysed by longitudinal logistic regression. Survival analysis included Cox regression and comparison to matched Australian population mortality rates (18 years of follow-up). ICD10-defined COD were analysed by competing risks regression.

Results: Four distinct pain/health status trajectory groups were identified, ranging from "better" to "poorer", each with relatively stable trajectories over time. Hospitalisation rates were highest in those with poorer pain/health status (OR 3.09, 95%CI 2.3, 4.2). Mortality was increased in participants with the poorer pain/health status (Group 4), either when compared to Groups 1-3 (HR 2.4, 95%CI 1.5, 3.8), or the Australian population (SMR 2.3, 95%CI 1.6, 3.3, $p<0.001$). Mortality was not increased in Groups 1-3 (SMR 1.1, 95%CI 0.8, 1.5). Circulatory diseases (cardiovascular and cerebrovascular) were the predominantly reported COD in this poorer pain/health status group compared to other trajectory groups (subhazard ratio 8.3, 95%CI 2.2, 31.7).

Conclusion: Poorer pain-related health status in RA is associated with higher hospitalisation and mortality rates, with circulatory diseases being the predominant COD.

Word count: 250

Pain is a highly prioritised symptom reported by patients with rheumatoid arthritis (RA)[1]. Persistent pain in RA may exist despite minimal disease activity and is often experienced as chronic widespread pain (CWP). Consequently, limited physical and social functioning, adverse psychological health and poor self-efficacy can occur and may contribute towards a downtrending health status trajectory[2, 3]. In a combined UK Biobank study and meta-analysis, in the general population, CWP was highly associated with increased mortality rates, primarily driven by poor lifestyle factors[4]. The burden of persistent pain in RA is well-recognised and mortality gap is globally higher in RA[5, 6]. However, the relationship between persistent pain and poor health outcomes in RA is largely unknown.

Using the Australian Rheumatology Association Database (ARAD), a large voluntary national registry with longitudinal self-reported data on treatment and health outcomes in inflammatory arthritis, we first identified distinct trajectory groups of pain-related health states in participants aged 25-75 years old and with rheumatologist-diagnosed RA. These participants entered ARAD within 5 years of diagnosis and with at least 3 years of follow-up and provided data linkage consent. Mortality data linkage to the Australian Death Registry was performed in early 2020. Distinct multi-trajectories using five different self-reported pain-related health outcome measures (Health Assessment Questionnaire Disability Index, visual analogue scores for pain, arthritis, global health, and the Assessment of Quality of Life utility index) were identified using latent discrete mixture modelling. Using these identified trajectory groups as predictors, we then examined their relationships with hospitalisations, mortality risk and causes of death (COD). Hospitalisation events were analysed by longitudinal logistic regression with covariates (gender, time-varying age, baseline comorbidity index and trajectory groups). Survival analysis included Cox regression stratified by gender and included covariates (diagnosis, age, and baseline comorbidity index). Mortality rates were matched to the Australian population mortality rates over 18 years of follow-up and standardised mortality ratios (SMRs) were calculated. ICD10-defined COD were analysed by competing risks regression.

From 806 included ARAD participants (72% female, mean age at diagnosis 54 years), four distinct pain/health status trajectory groups were identified, ranging from “better” to “poorer”, each with relatively stable trajectories over time (Figure S1). Participants with the poorer pain/health status (Group 4) were more likely to be obese and have a higher comorbidity index and lower socioeconomic status (Table 1). Hospitalisation rates were highest in those with poorer pain/health status (odds ratio, OR 3.09, 95%CI 2.3, 4.2) (Table S1). Excess mortality was observed in participants with the poorer pain/health status (Group 4), either when compared to Groups 1-3 (hazard ratio, HR 2.4, 95%CI 1.5, 3.8), or the Australian population (SMR 2.3, 95%CI 1.6, 3.3, p<0.001) (Table 2; Table S2 and Figure S2). Notably, mortality was not increased in Groups 1-3 (SMR 1.1, 95%CI 0.8, 1.5). Circulatory diseases (cardiovascular and cerebrovascular) were the most frequently reported COD in this poorer pain/health status group compared to other trajectory groups (subhazard ratio, SHR 8.3, 95%CI 2.2, 31.7) (Tables S3 and S4).

This is the first study to demonstrate the association between poorer pain-related health status in RA and higher hospitalisation events and excess mortality, with circulatory diseases being the predominant COD. More importantly, this study has suggested that the mortality gap remains significantly higher in those with RA and poorer pain-related health status compared to the Australian general population. Overall, the burden of increased mortality risk in RA may be largely borne by this identifiable subset of patients, implying a need to implement timely and strategic multi-dimensional pain management and in parallel, to consider the overall health and well-being of these patients throughout their disease trajectory.

Word count: 594 (limit 600)

Acknowledgement

The authors sincerely acknowledge and thank the rheumatologists and ARAD participants for their ARAD data contribution. The authors would like to thank the ARAD Steering Committee and the ARAD Database staff members.

Ethics Approval

All ARAD participants provide written permission to be contacted by ARAD investigators and written informed consent for study participation and for anonymous data analyses and associated data linkages. Ethics approval for ARAD has been granted by 18 committees and organisations across all Australian states and territories.

Patients and Public Involvement

Patients or the public were directly not involved in the design, or conduct, or reporting, or dissemination plans of this specific study. We acknowledge the patients for their ARAD data contribution.

Author Contributions

HLP, SL, CLH and SLW contributed to the conception and design of the work, the acquisition, analysis and interpretation of the data. HLP contributed to the main drafting and the writing of the manuscript and all authors (HLP, SL, OR, RJB, JT, BR, CB, ML, LM, RB, CLH, SLW) contributed to the subsequent drafting and revision of the manuscript. All authors (HLP, SL, OR, RJB, JT, BR, CB, ML, LM, RB, CLH, SLW) read and approved the final manuscript.

Data Availability Statement

The dataset used and analysed specific for this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors have declared no conflicts of interest.

Funding

The Australian Rheumatology Association Database (ARAD) receives support of unrestricted educational grants administered through the Australian Rheumatology Association from Pfizer Australia, AbbVie Pty Ltd, Eli Lilly Australia Pty Ltd, Sanofi Australia, Celgene Australian & NZ, Bristol Myers Squibb Australia Pty Ltd, Amgen Australia Pty Ltd, Aventis, AstraZeneca, and Roche. ARAD was previously supported by an Australian National Health and Medical Research Council (NHMRC) Enabling Grant (ID 384330). Infrastructure support for ARAD is provided by Cabrini Health, Monash University, Royal North Shore Hospital, and the Australian Rheumatology Association (ARA).

This manuscript formed part of HLP's Ph.D. work, for which HLP was the recipient of the Ken Muirden Overseas Training Fellowship from the Arthritis Australia, an educational grant funded by the ARA, and HLP also received the Australia Postgraduate Scholarship from the University of Adelaide. RB is funded by an NHMRC Investigator Fellowship.

References

1. Geenen R, Overman CL, Christensen R, Asenlof P, Capela S, Huisinga KL, et al. EULAR recommendations for the health professional's approach to pain management in inflammatory arthritis and osteoarthritis. Ann Rheum Dis. 2018 Jun;77(6):797-807. DOI: 10.1136/annrheumdis-2017-212662. Cited in: Medline.
2. Lee YC, Cui J, Lu B, Frits ML, Iannaccone CK, Shadick NA, et al. Pain persists in DAS28 rheumatoid arthritis remission but not in ACR/EULAR remission: a longitudinal observational study. Arthritis Res Ther. 2011 Jun 8;13(3):R83. DOI: 10.1186/ar3353.
3. Walsh DA, McWilliams DF. Mechanisms, impact and management of pain in rheumatoid arthritis. Nat Rev Rheumatol. 2014 Oct;10(10):581-92. DOI: 10.1038/nrrheum.2014.64.
4. Macfarlane GJ, Barnish MS, Jones GT. Persons with chronic widespread pain experience excess mortality: longitudinal results from UK Biobank and meta-analysis. Ann Rheum Dis. 2017 Nov;76(11):1815-22. DOI: 10.1136/annrheumdis-2017-211476. Cited in: Medline.
5. Dadoun S, Zeboulon-Ktorza N, Combescure C, Elhai M, Rozenberg S, Gossec L, et al. Mortality in rheumatoid arthritis over the last fifty years: systematic review and meta-analysis. Joint Bone Spine. 2013 Jan;80(1):29-33. DOI: 10.1016/j.jbspin.2012.02.005. Cited in: Medline.
6. Finckh A, Gilbert B, Hodkinson B, Bae SC, Thomas R, Deane KD, et al. Global epidemiology of rheumatoid arthritis. Nat Rev Rheumatol. 2022 Oct;18(10):591-602. DOI: 10.1038/s41584-022-00827-y. Cited in: Medline.

TABLES

Table 1: Baseline data on general demographics, socioeconomic demographics, comorbidity index, pain-related health outcome measures and medication use for participants stratified by pain-related health status trajectory groups. Trend tests (p_{trend}) were performed using the Jonckheere-Terpstra test

Table 2: Standardised Mortality Rates (SMR) compared to age, sex and calendar year matched Australian population mortality rates

Table 1: Baseline data on general demographics, socioeconomic demographics, comorbidity index, pain-related health outcome measures and medication use for participants stratified by pain-related health status trajectory groups. Trend tests (p_{trend}) were performed using the Jonckheere-Terpstra test

Baseline	Group1 (better)	Group2	Group3	Group4 (poorer)	p-val[1]	ALL
N	143 (18%)	232 (29%)	255 (32%)	176 (22%)		806
Age at Diagnosis: mean (sd)	51 (12)	51 (12)	54 (11)	54 (11)	0.022	52 (11)
Age at ARAD entry: mean (sd)	52 (12)	53 (12)	55 (11)	55 (11)	0.013	54 (11)
Disease duration: mean (sd)	2.4 (1.3)	2.1 (1.3)	2.4 (1.3)	2.4 (1.4)	0.37	2.3 (1.3)
Females: n (%)	82 (57%)	159 (69%)	198 (78%)	139 (79%)	<0.001	578 (72%)
BMI (WHO category)						
N	*105*	*173*	*171*	*105*		*554*
Normal	35 (33%)	52 (30%)	36 (21%)	12 (11%)	<0.001	135 (24%)
Overweight	51 (49%)	60 (35%)	52 (30%)	29 (28%)		192 (35%)
Obese	19 (18%)	61 (35%)	83 (49%)	64 (61%)		227 (41%)
Current smoker	12 (8%)	31 (13%)	48 (19%)	31 (18%)	0.026	122 (15%)
Education (N = 987)						
Did not complete High School	33 (23%)	57 (25%)	76 (30%)	62(35%)	0.001	228 (28%)
Completed High School	46 (32%)	56 (24%)	83 (33%)	53 (30%)		238 (30%)
Post High School	64 (45%)	119 (51%)	96 (38%)	61 (35%)		340 (42%)
Comorbidity Index: mean (sd)	0.6 (0.8)	1.0 (1.2)	1.3 (1.4)	2.0 (1.6)	<0.001	1.2 (1.4)
SES quintile[2] (SA1), : n (%)						
N	*119*	*186*	*221*	*156*		*682*
1 (lowest)	14 (12%)	38 (20%)	50 (23%)	39 (25%)	<0.001	141 (21%)
2	29 (24%)	35 (19%)	43 (19%)	40 (26%)		147 (22%)
3	28 (24%)	30 (16%)	55 (25%)	37 (24%)		150 (22%)
4	18 (15%)	44 (24%)	44 (20%)	23 (15%)		129 (19%)
5 (Highest)	30 (25%)	39 (21%)	29 (13%)	17 (11%)		115 (17%)
HAQ-DI[3]: mean (sd)	0.31 (0.46)	0.69 (0.54)	1.19 (0.50)	1.70 (0.50)	<0.001	1.00 (0.70)
Pain VAS[4]: mean (sd)	24.6 (22.9)	39.9 (23.1)	50.7 (21.7)	62.8 (21.1)	<0.001	45.6 (25.5)
Arthritis VAS score: mean (sd)	25.5 (24.6)	40.6 (24.5)	52.1 (22.5)	61.9 (21.7)	<0.001	46.1 (26.3)
Patient global health: mean (sd)	78.0 (15.8)	69.3 (16.7)	59.7 (17.2)	49.1 (18.8)	<0.001	63.5 (19.8)
AQoL[5] utility index: mean (sd)	0.75 (0.18)	0.63 (0.17)	0.48 (0.18)	0.27 (0.17)	<0.001	0.53 (0.24)
Any DMARD[6]: n (%)	136 (95%)	213 (92%)	220 (86%)	155 (88%)	0.001	724 (90%)
Methotrexate: n (%)	126 (88%)	196 (84%)	186 (73%)	139 (79%)	0.003	647 (80%)
bDMARDs[7]: n (%)	81 (57%)	122 (53%)	144 (56%)	110 (63%)	0.16	457 (57%)
Prednisolone: n (%)	58 (41%)	126 (54%)	131 (51%)	103 (59%)	0.008	418 (52%)
Opioids: n (%)	24 (17%)	65 (28%)	105 (41%)	102 (58%)	<0.001	296 (37%)
NSAIDs[8]: n (%)	64 (45%)	121 (52%)	130 (51%)	79 (45%)	0.85	394 (49%)

[1]Jonckheere-Terpstra non-parametric test for ordered alternatives
[2]Socioeconomic Status (SES) was measured by the Index of Relative Socio-economic Advantage and Disadvantage (IRSAD)
[3]Health Assessment Questionnaire Disability Index
[4]Visual Analogue Score
[5]Assessment of Quality of Life
[6]Disease-modifying anti-rheumatic drug

[7]*Biologic DMARDs*
[8]*Non-steroidal anti-inflammatory drugs*

Table 2: Standardised Mortality Rates (SMR) compared to age, sex and calendar year matched Australian population mortality rates

Trajectory Group	N	Observed deaths	Expected deaths	person-years	SMR (95% CI)	p-val
Group 1,2,3	630	45	40.41	5624.32	1.11 (0.83, 1.49)	0.47
Group 4	176	29	12.50	1493.94	2.32 (1.61, 3.34)	<0.001
All	806	74	52.91	7118.26	1.40 (1.11, 1.76)	0.004

SUPPLEMENTARY FILE

Figure S1: Longitudinal trajectories for each of the defined participant trajectory subgroups for each of the five outcomes: Health Assessment Questionnaire Disability Index (HAQ-DI), Pain visual analogue score (VAS), Arthritis condition VAS, patient global health assessment and Assessment of Quality of Life (AQoL) utility score

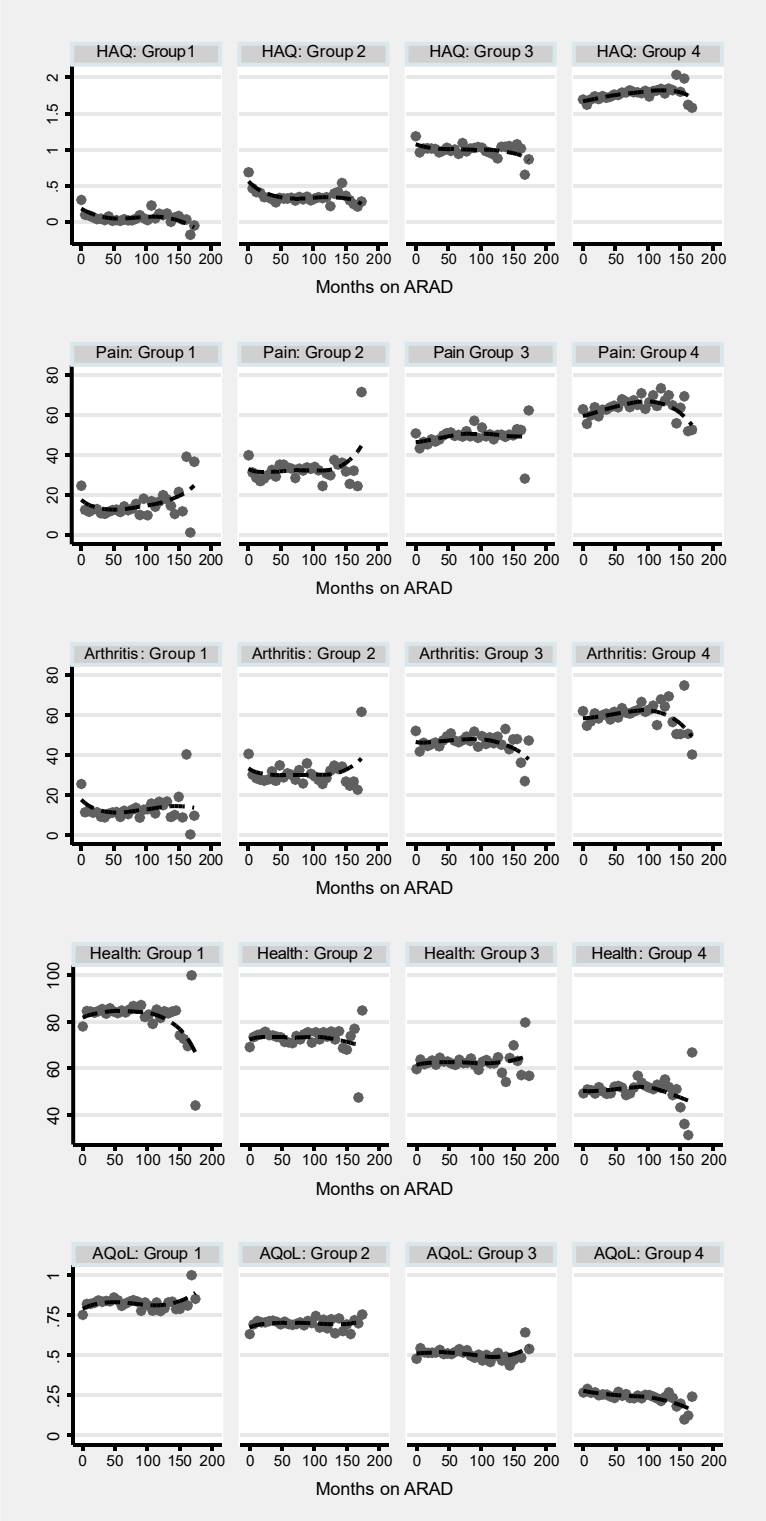


Figure S2: Kaplan-Meier Survival curves for participants stratified by pain-related health status trajectory groups

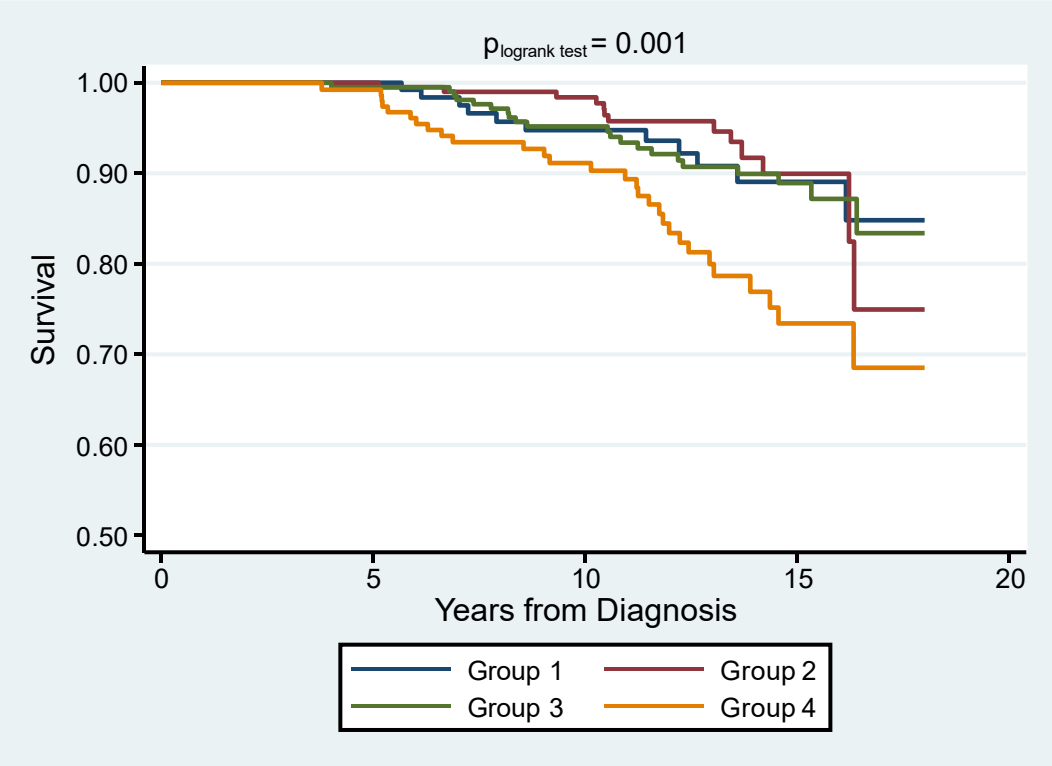


Table S1: Longitudinal (panel) regression logistic regression model for hospitalisation[1]

Covariate	Odds Ratio (95% CI)	p-value
Trajectory Group:		
Group 1	1 (base)	
Group 2	1.30 (0.99, 1.72)	0.059
Group 3	1.87 (1.42, 2.45)	<0.001
Group 4	3.09 (2.29, 4.17)	<0.001
Sex:		
Males	1 (base)	
Female	0.81 (0.67, 0.99)	0.039
Age	1.01 (1, 1.02)	0.019
Baseline comorbidity index	1.04 (0.98, 1.11)	0.20

[1]*Participants were asked if they had been in hospital since their last ARAD questionnaire*

Table S2: Cox proportional hazards regression (stratified by sex), adjusted for age at diagnosis and baseline comorbidity index. Results were reported as Hazard Ratios (HR) with 95% confidence intervals (95% CI)

Covariate	HR (95% CI)	p-val
Trajectory Group		
Group 1	1 (base)	
Group 2	0.96 (0.42, 2.19)	0.93
Group 3	1.24 (0.58, 2.65)	0.58
Group 4	2.58 (1.23, 5.39)	0.012
Age at diagnosis	1.10 (1.07, 1.08)	<0.001
Baseline comorbidity index	0.91 (0.78, 1.08)	0.28

Table S3: Underlying cause of death (COD) by ICD10 Chapter. Results are reported as SubHazard Ratios (SHR), with 95% confidence intervals (95% CI), from a competing risks regression analysis for Trajectory Group 4 compared to Groups 1,2,3 combined

COD ICD Chapter	ICD10 codes	Group 1,2,3	Group 4	Total	SHR (95% CI)	p-val
N		630	176	806		
Cancer	C00.0/D48.9	13	8	21	2.20 (0.92, 5.30)	0.077
Circulatory	I00.0/I99.9	3	7	10	8.26 (2.16, 31.69)	0.002
Respiratory	J00.0/J99.9	3	2	5	2.38 (0.40, 14.31)	0.34
Musculoskeletal	M00.0/M99.9	3	4	7	4.77 (1.08, 21.13)	0.040
Other		11	5	16	1.63 (0.60, 4.68)	0.36
Not reported		12	3	15		
All deaths		45	29	74		

Table S4: Itemised Causes of Death

CAUSE OF DEATH	N
A00–B99: Certain infectious and parasitic diseases	
A09.9 Gastroenteritis and colitis of unspecified origin	1
C00–D48: Neoplasms	
C15.9 Malignant neoplasm: Oesophagus, unspecified	1
C16.9 Malignant neoplasm: Stomach, unspecified	1
C19 Malignant neoplasm of rectosigmoid junction	2
C20 Malignant neoplasm of rectum	1
C25.9 Malignant neoplasm: Pancreas, unspecified	1
C26.0 Malignant neoplasm: Intestinal tract, part unspecified	1
C32.0 Malignant neoplasm: Glottis	1
C34.9 Malignant neoplasm: Bronchus or lung, unspecified	5
C43.9 Malignant neoplasm: Malignant melanoma of skin, unspecified	1
C45.9 Mesothelioma, unspecified	2
C49.9 Malignant neoplasm: Connective and soft tissue, unspecified	1
C50.9 Malignant neoplasm: Breast, unspecified	2
C71.9 Malignant neoplasm: Brain, unspecified	1
D43.2 Neoplasm of uncertain or unknown behaviour: Brain, unspecified	1
E00–E90: Endocrine, nutritional and metabolic diseases	
E11.7 Type 2 diabetes mellitus: With multiple complications	1
E11.9 Type 2 diabetes mellitus: Without complications	1
E14.2 Unspecified diabetes mellitus: With renal complications	1
G00–G99: Diseases of the nervous system	
G12.9 Spinal muscular atrophy, unspecified	1
G37.8 Other specified demyelinating diseases of central nervous system	1
I00–I99: Diseases of the circulatory system	
I21.9 Acute myocardial infarction, unspecified	1
I25.9 Chronic ischaemic heart disease, unspecified	3
I35.0 Aortic (valve) stenosis	1
I51.6 Cardiovascular disease, unspecified	1
I61.9 Intracerebral haemorrhage, unspecified	2
I63.9 Cerebral infarction, unspecified	1
I64 Stroke, not specified as haemorrhage or infarction	1
J00–J99: Diseases of the respiratory system	
J18.9 Pneumonia, unspecified	1
J44.0 Chronic obstructive pulmonary disease with acute lower respiratory infection	1
J47 Bronchiectasis	1
J84.9 Interstitial pulmonary disease, unspecified	1
J86.0 Pyothorax with fistula	1
K00–K93: Diseases of the digestive system	

CAUSE OF DEATH	N
K57.8 Diverticular disease of intestine, part unspecified, with perforation and abscess	1
K59.8 Other specified functional intestinal disorders	1
K81.9 Cholecystitis, unspecified	1
L00–L99: Diseases of the skin and subcutaneous tissue	
L03.9 Cellulitis, unspecified	1
M00–M99: Diseases of the musculoskeletal system and connective tissue	
M06.9 Rheumatoid arthritis, unspecified	5
M32.1 Systemic lupus erythematosus with organ or system involvement	1
M35.9 Systemic involvement of connective tissue, unspecified	1
N00–N99: Diseases of the genitourinary system	
N17.0 Acute renal failure with tubular necrosis	1
N18.5 Chronic kidney disease, stage 5	1
V01–Y98: External causes of morbidity and mortality	
V47.5 Car occupant injured in collision with fixed or stationary object: Driver injured in traffic accident	1
W19 Unspecified fall	1
X44 Accidental poisoning by and exposure to other and unspecified drugs, medicaments and biological substances	1
X64.5. Intentional self-poisoning by and exposure to other and unspecified drugs, medicaments and biological substances	1
COD Not reported	15
ALL DEATHS	74

7.4 Summary

Individuals with chronic pain have an increased mortality risk, primarily due to poor lifestyle risk factors, as identified in a large UK Biobank registry data. In this study using the ARAD dataset, using similar multi-trajectory analytic method presented in Chapter 6, four distinct trajectories of varying pain-related health status in individuals with RA, ranging from 'better' to 'poorer', were identified and their relationships with hospitalisation and mortality rates.

Main findings include:

- Increased hospitalisation rates were seen in those with poorer pain-related health status in RA
- In parallel, higher mortality risk was identified in those with poorer pain-related health status, when compared to the other trajectory subgroups and the Australian population. Of note, mortality risk in RA was not identified in the other trajectory subgroups
- In the trajectory subgroup of poorer pain-related health status, circulatory diseases (cardiovascular and cerebrovascular) were the most common reported causes of death

The above study findings were largely borne by this identifiable subset of patients, further stressing the importance of early identification of such patients with high pain and poor health status at the time of diagnosis of RA, and to allow appropriate interventions early on.

Summary of Section 3

Section 3 presents the results of chapters 5, 6 and 7, which address the third research question in this thesis in terms of evaluating the longitudinal impact of persistent pain in arthritis, and specifically in RA.

Stemming from the knowledge gained form the work presented in Chapter 4 on longitudinal data analysis, studies presented in this Section 3 have allowed me the opportunities to examine the relationships between pain trajectory and important pain-related health outcome measures in RA using Australian datasets.

In detail, the study presented in chapter 5 highlights the value of using of DAS28-P index score as a discriminatory measure of response to DMARD therapy in an early RA cohort.

In chapters 6 and 7, using multi-trajectory longitudinal data analysis on five different pain-related health outcome measures obtained from the ARAD dataset on individuals with established RA, study results have shown significant associations with some sociodemographic indicators (female predominance, obesity, low SES), medication use (opioids, prednisolone, NSAIDs), and higher hospitalisation and mortality rates in those with poorer pain-related health status. These key study findings have shed some light on the attributable burden of persistent pain in RA on treatment and health outcomes, of which such impact may be reversible or prevented if these at-risk individuals with high pain and poor health status are identified early on at the time of diagnosis.

Chapter 8: Discussion

8.1 Preface

This chapter presents the overall synthesis of the evidence that emerge from the work completed during my PhD candidature.

First, I present an overview of all the research work presented in this thesis, consisting of the emerging themes followed by detailed discussions of the key findings for each theme. Second, I present the significance and implications of these key findings in real-world clinical practice and research. Third, as my remote candidature involved a significant amount of time spent overseas, my experience as an overseas postgraduate student and the development of my research skills, both remotely in the UK and locally in Australia, are presented. Fourth, as Covid-19 hit mid-part during my PhD candidature, I present some of the minor Covid-19-related challenges faced. Lastly, I end this chapter by discussing the strengths and limitations of my research work.

8.2 Overview of the Research

Overall, this thesis explores the potentials of longitudinal characterisations of pain in RMD. This thesis has provided the evidence of the benefits and challenges of using digitalised health-related data to inform patterns of pain variability in RMDs using intensive longitudinal methods and the possibility of capturing at-risk individuals having persistent pain in RA and important negative corollaries that followed in terms of medication use and health-related outcomes.

Chronic pain in RMD is prevalent worldwide. At an individual level, the management of chronic pain in arthritis is multifaceted, and within the body of research in this thesis, three emerging themes are identified:

- Mechanisms of pain measures
- Longitudinal pain patterns in RMDs
- Longitudinal effects of persistent pain in RA

Next, I will discuss the abovementioned themes and their corresponding key findings.

8.3 Key Research Findings, Significance and Implications for Clinical Practice and Research

8.3.1 Mechanisms of pain measures

The interpretation of the pain scales and the frequency of pain measurements used matter in characterising pain longitudinally

When patients are asked about their pain level, often, VAS and NRS of pain level are commonly used in describing their pain in arthritis and this is most often done retrospectively. In RA, based on the systematic scoping review presented in this thesis, a heterogenous pattern of study designs and outcomes is seen, in terms of the types of pain scales used as well as the number of timepoints used to longitudinally characterise pain patterns over time. Such heterogeneity in study outcomes limits the comparison between study results, especially across clinical trials and observational studies. In addition, VAS or NRS of pain level are usually considered as ordinal scales, which assumes similar interval of difference between each pain level in these scales in order, although the magnitude of this interval of difference is not constant, as seen in a temperature range[209].

As we know that pain perception is subjective to individuals and the lived experience of chronic pain in arthritis can fluctuate with time, the magnitude of such measure of pain level in an ordinal scale may not be the same from time to time, even in the same individual. In general, even though VAS or NRS are easily applied in most history taking of pain symptoms and is assumed to measure a unidimensional component of pain, caution is required when interpreting these pain scales. As shown in Chapter 5 of this thesis, the DAS28-P index clearly identifies at-risk individuals with non-inflammatory pain in early RA, based on the persistently higher level of the subjective component of VAS PGA within the composite score of disease activity, DAS28, regardless of the frequency of timepoints of disease outcome measures.

Predictors of pain influence pain reporting – an important consideration in longitudinal studies

In the scoping review of this thesis, it is recognised that gender, sociodemographic and socioeconomic indicators, psychological status, and sleep are important predictors that influence pain level reporting and pain fluctuation. Similarly, as demonstrated in the studies presented in Chapters 5, 6, and 7, these important predictors, along with other pain-related health outcomes, are also instrumental in explaining the time-varying effects on medication use and health outcomes, especially in those with persistent pain. One of the main features seen as either a predictor or confounder in the scoping review presented in this thesis is female predominance. For instance, higher tendency of pain fluctuation and higher pain level reporting were seen in female cohort[187, 203]. Similarly, female predominance is a prominent feature in most of my studies done in RA cohort presented in this thesis, partly highlighting the well-established natural history of RA, that affects mostly women with a female to male ratio of 3:1[212, 213]. Gender disparity influences the reporting of PROs and pain experience, and as such, for trials of evaluating pain and treatment response in RA, female predominance is an important confounder to consider in the interpretation of the study results of trials evaluating pain and treatment response in RA[214, 215]. Cross-cultural differences in pain reporting inherently exists in clinical practice and research, therefore, an important factor to consider when individuals are asked to express their pain level[211].

Changes in pain over time in arthritis is minimal, regardless of the frequency of pain measurements

The scoping review in this thesis has identified stable pain trajectories in RA, regardless of the frequency of timepoints of pain measures. Similarly, stable pain trajectories in OA were identified in another systematic review by Previtali and colleagues[227]. These findings highlight the advantages of using trajectory analytical methods in evaluating the inter-individual variation of pain over time and identifying at-risk individuals such as those with persistent high level of pain or those with increasingly worse pain. However, such analytical

approach has limits in clinical practice when concerns of flares or response shift in pain level cannot be discerned from an individual level. In Chapter 4 of this thesis, from the *Cloudy with a Chance of Pain* study, patterns of pain improvement are identified in those with higher initial pain scores across different inflammatory and non-inflammatory RMDs, potentially suggesting regression of mean. It is important to note that in this study, a subset cohort data over the first 30-day period was chosen for the analysis, and therefore, highlights the importance of considering this phenomenon in repeated data when random errors exist[263]. Such random errors in terms of within-individual variation in pain may be instrumental in explaining the magnitude of change in pain, which is not explicitly described in the trajectory analysis.

The role of patient global assessment within the DAS28 composite score remains contentious

DAS28 is commonly used in clinical practice and research to monitor disease activity and as a proxy for guiding treatment decision. The study in Chapter 5 has shown not only the potentials of using DAS28-P index as a discriminatory measure in treatment response in early RA, but the implicit concerns on relying upon the use of DAS28 scores in the current treat-to-target approach in achieving disease remission in RA. Specifically, the role of the PGA within the DAS28 remains contentious. A large multinational study using the METEOR database has shown evidence of high level of PGA in patients having disease remission in RA, with implications of predisposing these patients with unnecessary immunosuppressive effect from escalation of DMARD therapy[281]. However, even though PGA is not predictive of radiographic progression based on a large individual patient data meta-analysis study, it is shown that PGA remains instrumental in defining patient's view of disease remission. It is also important to note that PGA has been demonstrated to be a limiting factor in defining disease remission in RA. For instance, as shown in a study by Studenic and colleagues, two-thirds of their RA outpatient cohort achieved 'near-remission', defined as the fulfilment of three of four Boolean-based criteria, with PGA being the primary limiting variable in

defining complete remission[282]. As such, a curated approach of incorporating PGA within the measures of disease activity is required[282].

8.3.2 Longitudinal pain patterns in RMDs

Digitalised patient-generated health-related data enables longitudinal characterisation of symptoms

The study presented in Chapter 4 using the Cloudy with a Chance of Pain data subset has shown the potentials of using pain symptom data obtained from a smartphone app to inform longitudinal patterns of pain over time. It is important to note the limitation of my study presented in Chapter in terms of using the first 30-day study period for data completeness. Factors influencing study engagement and attrition in mobile health studies should be considered when interpreting the study results[251, 283].

Pain trajectory analysis in arthritis allows stratification of individuals with varying degree of change in pain over time

As demonstrated in studies presented in Chapters 5, 6, and 7 in this thesis, distinct trajectory subgroups of individuals with varying degree of pain over time in both early and established RA are identified using the clustering methods of composite disease activity score and pain-related health outcomes respectively. As discussed earlier, at-risk individuals can be identified and tailored management of pain and appropriate modification of DMARD therapy can be delivered. By stratifying individuals with varying patterns of pain over time, an overview of different pain phenotypes that exist in RA is captured, further confirming the nature of inter-individual variability of pain in individuals living with RA, despite differences in psychosocial factors, sociodemographic indicators, treatment, and comorbidities.

The magnitude of changes in pain over time is not fully captured in traditional pain trajectory analysis

As discussed earlier, there are limitations in terms of the usefulness of trajectory subgroups in informing intra-individual changes in pain, and specifically, in dissecting the magnitude of change in pain over time. The response shift of pain is common in individuals experiencing chronic pain, whereby the reporting of pain can differ over time due to adaptation of pain response over time, even if the pain level experienced may not differ as much[284, 285]. Additionally, substantial intra-individual variability in the dynamic process of pain experience is not captured in traditional pain trajectory analysis, therefore, limiting the granularity behind our understanding of the influence of pain variability on pain processing, coping, and treatment response[219]. Alternative methods of assessing the magnitude of pain variability in real-time is required.

8.3.3 Longitudinal effects of persistent pain in RA

Pain in RA is much more than just a symptom related to the arthritis itself

In the multi-trajectory analysis of pain-related health outcomes in studies presented in Chapters 6 and 7 in this thesis, pain in RA is no longer viewed as a physical symptom related to the disease itself. In fact, the pain experience and the health status of the individuals are interdependent. In detail, the major difference between the trajectory groups is readily identifiable as the location (level) of the scores for each pain-related health outcome, rather than the shape of the trajectories over time, and in fact, changes over time within each group are relatively small. More importantly, the trajectory patterns across these five different pain-related health outcomes, when compared individually, are remarkably similar, indicating that they measure the same underlying (latent) pain-related health construct. Such findings extend the understanding that the management of pain requires a

holistic and multidimensional approach of evaluating the overall well-being of individuals, especially those at-risk individuals with poorer pain-related health status trajectory.

Disproportionate adjustment in dose and type of DMARD therapy in RA is seen in individuals with non-inflammatory pain and poorer pain-related health status

As shown in studies presented in Chapters 5 and 6 in this thesis, a higher tendency to either modify DMARD therapy or to increase treatment dose unnecessarily is seen in those with non-inflammatory pain and similarly, in those with poorer pain-related health status. As shown in study in Chapter 5, uptitration in DMARD dosage, including glucocorticoids, is seen in those with non-inflammatory pain subgroups, which is primarily driven by the subjective component scores within the DAS28. Similarly, as shown in study in Chapter 6, those with poorer pain-related health status are more likely to use glucocorticoids and non-TNF inhibitor biologic therapy, and to experience more b/tsDMARD treatment modifications, consistent with more refractory and difficult-to-treat disease. These at-risk individuals with RA are potentially predisposed to intermediate- and long-term DMARD-related and glucocorticoid-related adverse effects. In individuals with persistent pain and poorer health status, a mindful evaluation of treatment decision and disease-related management is required at the time of diagnosis and throughout the disease course, in parallel to their pain experience.

Opioid use in RA is at an alarming level for at-risk individuals with poorer pain-related health status trajectory

In the study presented in Chapter 6, opioid use in individuals with poorer pain-related health status is alarmingly high, which highlights a consistent finding with a prior study of opioid use in the ARAD cohort which concluded that NSAID and DMARD treatment did not obviate opioid use in all patients[286]. Evidence for the benefits of either weak or strong opioid use in treating RA pain is insufficient, and as such, a conditional recommendation against opioid

use in the latest Australian Living Guideline for the treatment of inflammatory arthritis is proposed[287, 288]. Concerningly, a recent American study has demonstrated that despite increasing awareness of the risks and harms associated with opioid use, chronic opioid use approximately doubled in patients with RA between 2002-2015, and was associated with pain, anti-depressant use, high disease activity and disability[289]. Opioid use in these at-risk individuals should be minimised and long-term use should be avoided. Additionally, at-risk individuals with persistent pain and using opioids should be identified and intervened early in the disease course to avoid opioid-related harms.

An increase in mortality risk in RA is a major concern in individuals with persistent pain, especially of those with poorer health status trajectory

As demonstrated in study presented in Chapter 7, excess in mortality is increased in those with persistent pain and of those with poorer pain-related health status, along with having the highest hospitalisation rates. Circulatory diseases (cardiovascular and cerebrovascular) are the most common contributing causes of death in this at-risk cohort in this study. Such finding is a concern on how we currently manage these at-risk individuals, when the mortality gap in RA itself is globally higher[280, 290]. In a study investigating excess mortality in the same ARAD cohort, the mortality gap in RA increases with time[291]. It is imperative to consider, in parallel, the influence of persistent pain on health outcomes in RA and the impact of other well established risk factors that affect mortality in RA.

8.4 Research Skills Attainment

During my remote candidature period in the UK, as I was working with the Cloudy with a Chance of Pain dataset, I have developed my epidemiology knowledge and research skills incrementally, particularly with my statistical programming skills in R, Stata, and Python. I can independently write the statistical code scripts and interpret the analysis output. As a newly qualified rheumatologist at the start of my candidature period, this was a steep learning curve as a beginner in clinical research. Furthermore, the effort in preparing and analysing such a large dataset is a huge undertaking, only made possible with the privilege to work with a multidisciplinary team of different background (epidemiologists, rheumatologists, meteorologists, statisticians, mathematicians, postgraduate students with heath informatics background, project manager, and the patient and public engagement representatives).

Moving forward, as I gained my understanding of analysing pain trajectory and pain variability using the Cloudy with a Chance of Pain dataset, I could continue such work using local datasets in Australia when I returned to South Australia, namely the Early Arthritis Cohort in South Australia and the national ARAD registry.

Most of the chapters presented in this thesis have been prepared in the form of manuscripts for publications. Being able to prepare and draft my research work into manuscripts have allowed the development of my critical thinking and appraisals of literature reviews, as well as constructing sound feedback to co-authors' comments.

8.5 Covid-19 Impact on the Research

In August 2022, towards the final part of my PhD candidature, I was confirmed Covid-19 positive based on a polymerase chain reaction (PCR) test done following the known close contact exposure of positive Covid-19 case in one my family member. Consequently, I had to take 2 weeks of time off research as I was symptomatic and was the primary carer of a young child in the family who was also affected. I was able to resume work following the 2-week break. There was no interruption with the workflow or access for research facility during this period.

8.6 Strengths of the Research

The strengths of the body of research in this thesis are:

- The advantages of having research skills and opportunities gained early in the postgraduate study could be applied throughout the completion of this thesis – these include statistical programming skills, digital epidemiology, research skills in mobile health studies and the foundation in using intensive longitudinal methods
- A unique opportunity to harness different types of data sources to address the research questions, including the application of complex statistical methods to analyse different data sources, including the mobile health study and the large observational cohort studies

8.7 Limitations of the Research

There are several limitations in the research work presented in this thesis, which are:

- In the work done using *Cloudy with a Chance of Pain* smartphone study, a short study period was selected during the analysis for data completeness. Hence, results from this exploratory work may be subject to selection bias, as highly engaged study participants were more likely to contribute data continuously.
- In the work done using the Australian based observational studies, study participants with RA were the focus in this thesis. Therefore, results from this work may not be generalisable to individuals with other forms of RMDs, such as OA and FM.
- Overall, the datasets used in the research for this thesis did not include information on disease activity, limiting further analysis of the relationship of the disease activity and the study outcome of interest such as pain-related health status. Furthermore, these datasets consist of mostly self-reported health-related information, which may be subject to recall bias or reporting bias.

8.8 Summary

This chapter summarises the overall key findings identified from the research groundwork presented in this thesis. Overarchingly, potentials of using digitalised health data in providing real-world evidence are outlined, along with the possibility of stratifying individuals of varying pain trajectories across different RMDs. Specifically in RA, persistent pain in RA is a concern in the current treat-to-target approach and identifying these at-risk individuals is pertinent. More importantly, evidence on the attributable burden of persistent pain on medication use and important health outcomes including mortality risk is presented. Next, I complete this thesis by presenting the conclusion and future directions.

Chapter 9: Conclusion and Future Directions

9.1 Conclusion

The study findings in this thesis have shown significant potentials that arise from the application of intensive longitudinal methods in analysing pain in arthritis. Pain trajectory analysis allows the opportunity to identify at-risk individuals, particularly of those with persistent pain and similarly, of those with varying degree of pain-related health status. Changes in pain over time in arthritis is minimal, regardless of the frequency of pain measurements. Perhaps, in the traditional pain trajectory analysis, alternative methods in assessing pain volatility are likely required. The magnitude of the burden of persistent pain and poor health status in RA is significant, with concerning trends of higher tendency of unnecessary treatment use and modification and excess mortality risk.

Most importantly, in the early days of the diagnosis of arthritis and throughout the disease course, pain and health status should be mindfully evaluated in parallel, with aims to improve the outlook of the individuals. Greater awareness of the adverse health-related impacts of persistent non-inflammatory pain in at-risk individuals should be a discernible priority amongst all treating rheumatologists and other health care professionals.

9.2 Future Directions

Based on the evidence seen in studies presented in this thesis, I present the following recommendations relevant for clinical practice and research:

1. Multidimensional measures of pain should be considered in longitudinal studies.

Studies identified through the scoping review have identified the limits of using the VAS or NRS in measuring pain level over time, especially in measuring disease activity in inflammatory arthritis. Other causes of pain such as neuropathic pain or non-inflammatory pain in patients with inflammatory arthritis may be best measured using alternative modalities, such as the painDETECT questionnaires or Patient Experienced Symptom State (PESS)[90, 292, 293]. Given that biopsychosocial factors can influence the pain pattern over time, a prognostic risk score using common biopsychosocial variables has potential roles in predicting the development and the spread of chronic pain over time, as shown in a recent data-driven UK Biobank study[294]. Such pain prognostication may optimise personalised pain management and targeted cohort selection in research and clinical trials[294].

2. Alternative statistical methods are required to assess the magnitude of pain variability.

One of the study limitations in examining pain variability across different RMDs is the assumption of linearity of the effect of time on change in pain over time. Such assumption is not necessarily true for the true dynamic ebb and flow of pain experience and in the traditional pain trajectory analysis, there are limits to explore the magnitude of pain variability. Non-linearity in dynamic time series and the constructs of intraindividual pain variability may potentially provide new key insights in how we define long-term and short-term pain variability in real-time.

3. Consider dual measures of disease outcomes in terms of treatment remission and symptom remission in future 'treat-to-target' strategy in RA.

Non-inflammatory pain largely influences the current composite disease activity measure, the DAS28, attributed to the PGA component. Studies by Ferreira and colleagues have recommended the use of dual treat-to-target (dual T2T) strategy, as defined by an aspect of controlling disease inflammation (biological remission) and another aspect of controlling disease impact (symptom remission). Focusing on capturing target information on disease impact, PESS and seven items of Rheumatoid Arthritis Impact of Disease (RAID.7) are some of the promising PGA tools that can be conveniently used in clinical practice[292, 295, 296]. Such alternative strategy may further support the acknowledgement of patient's view of their disease activity and subsequent patient-reported satisfaction of being in control of their disease trajectory and management. Risk of unnecessary modification of DMARD therapy can be minimised using this alternative strategy.

4. Consider objective measures of non-inflammatory pain in RA research.

Studies presented in this thesis have highlighted the importance of identifying patients with non-inflammatory pain in RA. Alternative measures to objectively quantify non-inflammatory pain in RA are possible such as the use of indices of pain sensitisation, and this is likely advantageous in the recruitment and stratification of participants of clinical trials when assessing treatment efficacy[297, 298]. In a proof-of-concept study by Wohlfahrt and her colleagues, lower knee pressure pain thresholds (PPT) and conditioned pain modulation (CPM) were shown to be predictive of DAS28 in those with low-moderate disease activity (pre-DMARD) and with higher baseline disease activity (post-DMARD) respectively[299]. Apart from solely characterising pain in RA, by using these indices of pain sensitisation measures, alongside the application of standard disease activity composite measures, such

assessment may allow future personalised mechanism-specific pain interventions in RA, targeting those with PGA-near remission[300].

5. Incorporate a mindful well-being evaluation of patients early from the outset at diagnosis in RA with the view to the development of appropriate management strategies in addition to suppression of inflammatory disease.

Integrative health approaches are recommended. These measures include accessible psychological and social welfare support, timely interventions in physical activity and lifestyle factors, and in parallel, consider the management of comorbidities and related modifiable risk factors that may influence the overall outlook of the health status[301, 302].

Appendices

Appendix A: Data Cleaning Process for Baseline Questionnaires (*Cloudy with a Chance of Pain*)

CLOUDY WITH A CHANCE OF PAIN -
THE PROCESS IN GENERATING THE BASELINE DATA SET

Huai Leng (Jessica) Pisaniello (21-23 February 2020)

TABLE OF CONTENT

APPENDICES

SUMMARY

This document provides a general overview of my work in preparing the baseline data set during my fellowship in Manchester. I have been using R and RStudio for this work (on-site and offsite using WakeIT access and Incline). There were some manual work done (i.e., changes made manually in the spreadsheet) for the free text columns in the baseline data set for conditions and medications. The pain site columns have not been looked into in more detail, in terms of identifying single vs multi-site pain.

Section 1 outlines the old baseline data set that I used at the start of my fellowship. This data set was valuable to me in terms of learning how to use R. The columns for this old baseline data set were used consistently for naming the variable columns in the new data set. There was a coding error in naming these columns, identified in May-June 2019, and corrections were made accordingly, and the work for this is outlined in Section 5.

Section 2 outlines the Cloudy Data dictionary and its file source.

Section 3 outlines the work in cleaning the free text conditions.

Section 4 and 5 outlines the work in preparing the final baseline data set, and also the work in preparing the Table One and flow chart for Cloudy's main paper. You will find that I have separated the work done in Section 4 (before June 2019) and the work done post-identification of motif coding error (after June 2019).

In each section, you will be able to find the relevant R scripts used for each step and the saved files in the R drive.

1. OLD BASELINE DATA SET (PRIOR TO MARCH 2018)

An old baseline data set (coded) was previously generated by the Cloudy team prior to my arrival in March 2018.

Directory: R:/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA
File name: baselinedata-clean.csv [dated:12/01/2018]

In this data set, the free text for cond.other has been coded. I believe that the work was done by a medical student (?APEP) under the supervision of the Cloudy team.

I have previously used this baseline data set in my R programming learning and I used this data set and old motif data set to generate some graphs as part of my R exercise. This baseline data set was not used as the final baseline data set for the Cloudy main paper.

2. CLOUDY DATA DICTIONARY

Directory: R:/Projects/CloudyWithAChanceOfPain/analyses/Jess/Cloudy Data/Data
File name: Cloudy data dictionary.pdf

I use this as my reference for coding the baseline data set and creating the Table One and flow chart for the Cloudy main paper.

3. CLEANING THE CHRONIC PAIN CONDITIONS IN THE BASELINE DATA SET

At the start of my research fellowship for the first few months in Manchester, whilst learning how to code using R, I took the opportunity to look through the free text for 'cond.other'. I also looked at the coded baseline data set to ensure that the free text has been coded correctly. I have noted a few observations of which the free text and the coded variables were discordant.

Examples:

1. When a study participant reported having osteoarthritis in the free text but did not tick the option 'OA' (i.e., the coded variable), in the old baseline data set, this study participant had two coded conditions, which were 'OA' and 'cond.other'. In fact, this participant only had one condition, which is 'OA' and the 'cond.other' should be NA/0.
2. When a study participant reported having 'psoriatic arthritis' in the free text, the participant was noted to have two coded conditions, which were 'SpA' and 'cond.other'. In fact, this participant had psoriatic arthritis, which should not be labelled as 'SpA' as there is no additional information to say if the psoriatic arthritis is an axial disease or peripheral joint disease or both. The option for 'SpA' variable is specifically pertaining to having either ankylosing spondylitis or axial spondyloarthropathy. Therefore, ideally, this participant should have one condition, which is 'cond.artr.unsp' (free text: psoriatic arthritis – see below) and the 'cond.spa' should be NA/0.

These are just some examples that prompted me to clean the baseline data set, mainly on the free text for chronic pain conditions. This work was done manually in mid-2018 using the csv spreadsheet and this work was saved in the R drive. I have previously mentioned this work to the Cloudy team during one of those regular fortnightly meeting.

You can find the categorised free text for the chronic pain conditions in this file.

Directory: R:/Projects/CloudyWithAChanceOfPain/analyses/Jess/Cloudy Data/Data
File name: FINAL-categorised-data-jess-11822.csv [dated: 23/08/2018]

This categorised free text (conditions) data set was eventually merged with a partially cleaned baseline data set, which I will outline the process in the next section.

An explanation for the variable names for this csv file is provided in Appendix 1.

3.1 FREE TEXT CHRONIC PAIN CONDITIONS THAT HAVE BEEN CODED INTO THE PRE-EXISTING CONDITION VARIABLES

Conditions mentioned in the free text that were eligible to be coded in the pre-existing 'condition' variables are listed as follow:

Variable Name	Free Text Conditions Eligible for Coding in the Pre-Existing Condition Variables
con.ra	Palindromic rheumatoid arthritis
con.oa	Wear and tear, worn thumb joints, joint wear, degenerate bone disease
con.spa	Spondylitis, cervical spondylitis, sacroiilitis, cranial spondylitis, sacroiliac arthritis
con.artr.unsp	Juvenile idiopathic arthritis (JIA), palindromic rheumatism, Still disease, synovitis, seronegative arthritis/arthropathy/polyarthritis, inflammatory arthritis, mono-inflammatory arthritis, arthralgia, inflammatory arthralgia, polyarthralgia, psoriatic, psoriatic arthritis, rheumatic pain, enteropathic arthritis, lupus polyarthritis, psoriatic rheumatism, joint pain from coeliac disease
con.cwpfm	Chronic pain syndrome, chronic pain from injury, chronic pain, central pain syndrome
con.mig	Tension headache, familial hemiplegic migraine, cluster headache, vestibular migraine, vascular migraine, hemicrania continua

Subsequently, once the above free text conditions were coded accordingly to the pre-existing 'condition' variables, then I made their corresponding 'con.other' variable as 'NA/0' unless the participant reported additional free text condition that could not be coded into the pre-existing 'condition' variable.

3.2 OTHER CHANGES MADE TO THE CHRONIC PAIN CONDITIONS

There were some participants with duplicate entry for the baseline. Therefore, it was decided that participants with the earliest entry date (based on the 'start.date.utc' variable) for their baseline data were included and the other duplicates with later date were excluded.

For the OA data entry, I found a number of duplicates (most participants had two duplicates, but some had more than two duplicates). Therefore, the presence or absence of having the diagnosis of OA was decided based upon the highest total number of similar responses. If the total was even, the earliest entry date (based on the 'start.date.utc' variable) was chosen to confirm the OA diagnosis.

For participants who selected as having 'con.none' but also had other 'condition' variable(s) selected, I changed this 'con.none' to NA/0.

4. PREPARING THE FINAL BASELINE DATA SET (PRIOR TO JUNE 2019)

4.1 RAW BASELINE AND MOTIF DATA SETS

These are important raw baseline and motif data sets that I use for my work in preparing the final baseline data set (prior to June 2019):

Directories:

- R:/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA
- R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/DATA

File names:

- Cloudy Health Questionnaire Baseline-report.xlsx [dated: 29/05/2017] – raw baseline data
- Custom Content Cloudy Health Questionnaire Baseline-report.xlsx [dated: 29/05/2017] – raw baseline data
- Osteoarthritis Questionnaire-report.xlsx [dated: 29/05/2017] – raw OA data
- cloudy-25-05-20172121212_motif_segmentvalue.csv [dated: 26/05/2017] – raw motif data
- Test_User.csv [dated: 02/03/2017] – raw test userids

4.2 BASELINE DATA PREPARATION

First, I prepared the baseline data using the first 3 files listed above and the work done can be found in this R script:

Directory: R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/Scripts
File name: 071018_baseline-data-preparation.R

Here, I provide you a detailed explanation for this R script:

1. Within the file 'Cloudy Health Questionnaire Baseline-report.xlsx', I found some non-functional users, i.e., 7 blank users and 3 of 'xxxxx' users, which I manually removed from the spreadsheet. In addition, I removed the column KOALAP from the spreadsheet. Given the changes that I made, I saved these changes as a new file:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: NEW-Cloudy Health Questionnaire Baseline-report.csv

In the R script, this file is called 'bsl.cloudy'. I changed the column names accordingly, as follow:

```
"weird.string", "sex", "yob", "post.code", "site.head", "site.face", "site.mouth.jaws",
"site.neck.or.shoulder.pain", "site.back.pain", "site.stomach.or.abdominal.pain",
"site.hip.pain", "site.knee.pain", "site.hands", "site.feet", "site.multi", "site.all", "site.none",
"cond.ra", "cond.oa", "cond.spa", "cond.gou", "cond.artr.unsp", "cond.cwpfm", "cond.mig",
"cond.np", "cond.other", "cond.none", "drug.pcmol", "drug.nsaids", "drug.simple.anal",
```

"drug.weak.opiates", "drug.strong.opiates", "drug.for.np", "drug.other", "drug.none", "drug.ster", "drug.dmard", "drug.bdmard", "drug.other.dmard", "belief", "belief.rain", "belief.cold", "belief.hot", "belief.dbp", "belief.dt", "belief.other", "via", "via.other", "userid", "portalid", "region", "redirect_to", "start.date..utc.", "submit.date..utc.", "network.id"

Note: I retained these variable names from the old baseline data set. However, at that time, I was not aware of the coding error of 'site.none', which is a non-existent variable. This error was discovered in June 2019 as I was cleaning the 'medication' variables.

Then, in 'bsl.cloudy', I changed the responses to 1 or 0 (i.e. coded responses for each variable), except the free text 'condition' and 'drug' variables.

I saved this coded 'bsl.cloudy' in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: CODED-CloudyHealth-11836.csv

Thereafter, I removed the test users from this 'bsl.cloudy', for which I created 'bsl.cloudyfinal' in the R script and this file was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: FINAL-CloudyHealth-11800.csv

2. I created a new file for the purpose of this work for file 'Custom Content Cloudy Health Questionnaire Baseline-report.xlsx', which was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: NEW-Custom Content Cloudy Health Questionnaire Baseline-report.csv

In the R script, this file is called 'bsl.custom'. I changed the column names accordingly, similar to above. Then, in 'bsl.custom', I changed the responses to 1 or 0 (i.e. coded responses for each variable), except the free text 'condition' and 'drug' variables.

I saved this coded 'bsl.custom' in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: CODED-CustomContent-351.csv

Thereafter, I removed the test users from this 'bsl.custom', for which I created 'bsl.customfinal' in the R script and this file was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: FINAL-CustomContent-351.csv

3. I combined both 'FINAL-CloudyHealth-11800.csv' and 'FINAL-CustomContent-351.csv' as 'bslcombo' in the R script and this combined data set was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: CLEAN-baselinecombined-12151.csv

4. In this 'bslcombo' data set, I identified 273 duplicates, of which I separated this file in R script as 'bslcombo.dupl' and this data set was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: baselinecombined-duplicates-273.csv

The non-duplicates were separated out from the above, labelled as 'bslcombo.clean' in the R script and this data set was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: CLEANFINAL-baselinecombined-11878.csv

5. In this 'bslcombo.clean', I subset those participants who were within the study period (20-01-2016 until 20-04-2017), labelled as 'baselineclean' in the R script and this data set was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: STUDY-baselinecombined-11822.csv

There were 56 users who were outside of the study period (between 17-12-2015 until 19-01-2016), labelled as 'bsl.extra' in the R script and this data set was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: OUTSTUDY-baselinecombined-56.csv

6. Now, I review the OA data set. Within the 'Osteoarthritis Questionnaire-report.xslx' file, there were 3 non-identifiable userid and therefore, these were removed manually in the spreadsheet and this new data set was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: NEW-Osteoarthritis Questionnaire-report.csv

I removed test users from this file and identified duplicates, labelled as 'OAduplicate' in the R script and was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: OA-duplicates-1013.csv

The non-duplicate OA file, labelled as 'OA.notdup' in the R script, was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: OA-clean-2169.csv

Note: I also manually checked this file and one userid 12022 required manual change of response due to >2 duplicates found for this userid. The response was changed based on the earliest date of entry.

7. I merged clean OA data set to the baseline data set, i.e., 'bslstudy' and 'OA.notdup', and this file was labelled as 'bslstudy.OA' in the R script, and this was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: STUDY-baselinecombined-mergedwithOA-11822.csv

8. I merged the 'bslstudy.OA' data set with the free text cond.other file 'FINAL-categorised-data-jess-11822.csv', and this file was labelled as 'bslfinal' in the R script and was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: STUDY-complete-baseline-11822.csv

Note: I manually changed 'con.other' to 'cond.other' as these columns are similar and also cond.other free text manually to con.other, in other to retain the free text column in the data set. This file was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: STUDY-COMPLETE-set-baseline -11822.csv

Then, for the above data set, I changed the NA to 0, and realign the 'cond.other' column to match the other columns in the data set. This file was labelled as 'bslcomplete' in the R script and was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: STUDY-baselineclean-11822.csv

4.3 MOTIF DATA PREPARATION

This work was mainly done to get the right baseline data set, and has not involved any modification to the original motif data set.

Directory: R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/Scripts
File name: 071018_motif-data-preparation.R

Here, I provide you a detailed explanation for this R script:

1. I first retrieved the motif data "cloudy-25-05-20172121212_motif_segmentvalue.csv", which was labelled as 'motif.raw' in the R script. I first removed the test userids from this data set (file was not saved due to its large volume). Then, I subset those who were within the study period (20-01-2016 until 20-04-2017), labelled as 'motifclean' in the R script and this file was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: STUDY-motifclean-11989.csv

The file having those who were outside the study period with motif data, labelled as 'motif.extra' in the R script, was saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: OUTSTUDY-motif-90.csv

4.4 MERGING BASELINE-MOTIF

This work can be found in:

Directory: R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/Scripts
File name: 071018_baselinemotif-merged-data-preparation.R

In this R script, I merged both 'bslclean' (from file "STUDY-baselineclean-11822.csv") and 'motifclean' (from file "STUDY-motifclean-11989.csv") by finding those userids who had only either one. I found 11,024 participants with both baseline and motif, 798 participants with baseline only (no motif) and 965 with motif only (no baseline).

I saved the file for those 11,024 participants, labelled as 'baseline.mot' in the R script, in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: STUDYPERIOD-baselinedata-correctmotif-11024.csv

I saved the file for those 798 participants, labelled as 'bslonly' in the R script, in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA
File name: STUDYPERIOD-baselineonly-798.csv

For those 965 participants with motif data only, this data set was labelled as 'motifonly' in the R script, but was not saved due to its large volume.

These 798 participants with baseline only and 965 participants with motif only were discussed in one of the ‘Crunch Time Cloudy’ meetings (together with Jamie, Anna and Belay) and it was decided that we only used the data for those within the study period and with both baseline and motif data. These steps taken were also discussed and agreed upon in the meeting.

I had other scripts subsequent to these work with regards to preparing the Table One and flow chart. However, when we discovered the coding mistake with the motif data set in June 2019, my role in preparing the Table One and flow chart was on hold until Anna cleaned the motif data accordingly. This work will be discussed in the next Section. I have not presented these scripts in this document but they can be retrieved through this directory:

Directory: R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/Scripts

5. PREPARING THE BASELINE DATA SET FOR CLOUDY MAIN PAPER (AFTER JUNE 2019)

5.1 CORRECTION FOR BASELINE DATA SET COLUMNS

In May 2019, as part of my PhD work, as I was cleaning the free text for the medications, I realised that the baseline columns were labelled wrongly in the old baseline data set (which were used consistently in my baseline preparation work in Section 4). I informed Will and the Cloudy team about this, and also double checked this using the Cloudy Data Dictionary as well as the original questions and options in the uMotif app. There is a Dropbox folder which contained these original questions and options for each question in the uMotif app.

This led to my work in correcting the baseline columns, for which this R script can be found in:

Directory: R:/CloudyWithAChanceOfPain/analyses/Jess/INCLINE_Jess/PhD/Data
File name: 210519_script_correctionforbaseline.R

Here, I provide you a detailed explanation for this R script:

I went back to the original file "STUDY-baselinecombined-mergedwithOA-11822.csv", and in the R script, this file was labelled as 'bslstudy.OA'.

I corrected the column names for:

1) site.none – changed to 'cond.final.none'
2) cond.none – changed to 'drug.final.none'
3) drug.none – changed to 'drug.dmard.none'

Then, I merged the free text condition file ("FINAL-categorised-data-jess-11822.csv") to the above data set 'bslstudy.OA', which is then labelled as 'bslfinal' in the R script.

Within this 'bslfinal' data set, I tried to realign the variables thst start with 'cond.' with the correctly labelled variables that start with 'con.'. Additionally, I check for multiple conditions and had new column created for this purpose. I also corrected the 'cond.final.none' to reflect those who truly reported no chronic pain condition.

As I need to manually correct the free text for the medications, I saved this 'bslfinal' data set in:

Directory: R:/CloudyWithAChanceOfPain/analyses/Jess/INCLINE_Jess/PhD/Data
File name: 210519_cleanbaseline_correctcodedconditions_11822.csv

Note: In the R script, this directory was not labelled correctly as I had to move around my folder due to my remote work away from UoM.

After I have manually corrected the medications, I saved the changes in this new file:

Directory: R:/CloudyWithAChanceOfPain/analyses/Jess/INCLINE_Jess/PhD/Data
File name: 210519_cleanbaseline_correctcondition_correctdrug_11822.csv

Within the same R script, I opened the above file "210519_cleanbaseline_correctcondition_correctdrug_11822.csv" and labelled it as 'newbsl' in the script. I changed the drug column names in order to sum number of drugs taken.

I also have additional column to sum the number of chronic pain condition for each participant. This is because in some of the free text conditions, these were not necessarily pain condition.

When these changes were made to the 'newbsl' data set, I saved this file in:

Directory: R:/CloudyWithAChanceOfPain/analyses/Jess/INCLINE_Jess/PhD/Data
Directory: R:/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA

File name: 210519_final_correctedbaseline_11822.csv

The above baseline data set is the data set that I use for the npj Cloudy main paper and for my PhD work. This data set is also the one that was discussed with Yuanyuan and Belay and the details for the variables can be found in this R markdown file, saved in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/Discussion_Jan2020
File name: 20200122_CloudyBaseline_DataPreparation.html [the R markdown code is embedded within this file, which can be found at the top right hand corner]

Details in how I cleaned the motif data and baseline data (i.e., to obtain the daysinstudy, number of entry, time in study, age, gender), I have learned the codes from Belay and Anna and used them consistently with my baseline work throughout. You can find the snippets of these codes in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/Discussion_Jan2020
File name: 20200122_Cloudy_MotifandBaseline_DataPreparation.html [the R markdown code is embedded within this file, which can be found at the top right hand corner]

For the full script details, I will outline them in Section 5.3.

Appendix 2 outlines the description for each variable in this baseline data set.

5.2 SCRIPT PART 1 FOR THE TABLE ONE

This is an initial R script prepared for case-crossover (CCO) analysis for n=3309. I have not outlined the details here as this was not the final CCO number in the main paper. You can find the R script in:

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA/TABLEONE

File name: 20190703_script_TableOne.R

5.3 SCRIPT PART 2 FOR THE FINAL TABLE ONE AND FLOW CHART FOR THE NPJ CLOUDY MAIN PAPER

This is the R script for the npj manuscript preparation, with regards to preparing the Table One and flow chart. There were corresponding emails from Anna on 020719 and 250719 with the following information required for this work and the correct number of participants for each part of the flow chart:

- 13207 users downloaded the app over the 12-month period
- 10584 participants with complete baseline data and motif data and at least one pain entry
- Need to exclude the first 10 days (for flow chart only)
- 2658 included in the final CCO analysis (no longer 3309)

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA/TABLEONE/manuscript
File name: 20190730_tableone_flowscript.R

(Please note: there is a similar script saved in R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA/TABLEONE and with similar name. This is an incomplete script and for some reason, I didn't change the name of the R script, which should have been done retrospectively.)

Here, I provide you a detailed explanation for this R script:

R packages used: tidyverse, data.table, sp, lubridate, tableone, bit64

1. Baseline data set (labelled as 'baseline' in the R script) was retrieved from the R drive.

Directory: R:/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA
File name: 210519_final_correctedbaseline_11822.csv

(This file is same as the one saved in the INCLINE_Jess folder)

I selected the columns required for this Cloudy main paper work – sex, yob, post.code, conditions, beliefs, userid, day

2. Motif data set (labelled as 'motif' in the R script) was retrieved from the R drive.

Directory: R:/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA
File name: 2090701_motif-weather-bsl-rain-sex_794007-rows_10584_userids.csv

Then I cleaned this motif data set using the following code, with subsequent clean motif labelled as 'cleanmotif' in the R script.

```
cleanmotif <- motif %>%
```

```
  dplyr::select(1:3, 8) ## select all motif variable

cleanmotif[is.na(cleanmotif)]<-0
cleanmotif$ind<-rowSums(cleanmotif[,symptom.start.col:symptom.end.col])
cleanmotif(cleanmotif$ind!=0)

cleanmotif<-cleanmotif %>%
  group_by(UserId) %>%
  dplyr::mutate(YMD=as.Date(day,format="%Y-%m-%d"),
                first = dplyr::first(YMD),
                last = dplyr::last(YMD))

cleanmotif<-cleanmotif%>%
  mutate(daysinstudy = difftime(last,first,units="days")+1)

cleanmotif<-cleanmotif%>%
  group_by(UserId) %>%
  dplyr::mutate(
    nentry = dplyr::n_distinct(YMD))
setDT(cleanmotif)

colnames(cleanmotif)

## Step 2b: Change calander time to process time
cleanmotif <- cleanmotif %>%
  group_by(UserId) %>%
  mutate(Date=as.Date(YMD,format="%Y-%m-%d"),
         prev.entry_date = c(0, diff(Date)))
cleanmotif <- ungroup(cleanmotif)

cleanmotif<-cleanmotif %>%
  group_by(UserId) %>%
  mutate(time = cumsum(prev.entry_date))
setDT(cleanmotif)

colnames(cleanmotif)
```

3. Then I retrieved the baseline data for 10,584 participants with motif data, labelled as ‘bsl.all’ in the R script. Then, I subset the data for the first 10 days exclusion, labelled as ‘tendayexcluded’ data set in the R script (6214 participants).

4. Cluster data set (prepared by Belay using the HMM method) was retrieved from the R drive.

Directory:
R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA/TABLEONE
File name: datawithcluster_newdat_checking.csv

Note: cluster labels:

- 1 – moderate engagement
- 2 – low engagement
- 3 – tourists
- 4 – high engagement

I left-joined this cluster data set (labelled as 'cluster' in the R script) with the 'bsl.all' data set, with the subsequent data set labelled as 'bslfinal' in the R script.

5. I cleaned the above 'bslfinal' data set to obtain gender and age. I also combined both cond.other and cond.none. I changed the response for the belief columns to either 1 or 0. Please see the following code used for this work.

```
bslfinal$yob[!(bslfinal$yob %in% c(1900:2000))] <- NA

sum(is.na(bslfinal$yob)) #401

bslfinal$yrentry <- year(as.Date(bslfinal$day, format = "%d/%m/%Y"))
bslfinal$age <- as.numeric(bslfinal$yrentry) - as.numeric(bslfinal$yob)

# Step 5c: Combine cond.final.none and condsum.pain.other
bslfinal$cond.otherplusnone <- bslfinal$cond.final.other + bslfinal$cond.final.none

# Step 5d: Change belief and other subgroups to 0 and 1 too
bslfinal$belief.rain <- if_else(is.na(bslfinal$belief.rain), 0, 1)
bslfinal$belief.cold <- if_else(is.na(bslfinal$belief.cold), 0, 1)
bslfinal$belief.hot <- if_else(is.na(bslfinal$belief.hot), 0, 1)
bslfinal$belief.dbp <- if_else(is.na(bslfinal$belief.dbp), 0, 1)
bslfinal$belief.dt <- if_else(is.na(bslfinal$belief.dt), 0, 1)
bslfinal$belief.other <- if_else(is.na(bslfinal$belief.other), 0, 1)
```

I saved this 'bslfinal' in:

Directory: R:/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA
File name: 20190729_10584_baselineWITHcluster.csv

6. CCO motif data set was retrieved from the R drive (labelled as 'motif.cco' in the R script):

Directory: R:/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA
File name: 20190719_CCO-data_NatDigMedpaper_2658-userids-6431-strata_91422-rows.csv

Then, I retrieved the baseline data for these CCO participants from the 'bslfinal' data set. I saved this 'bslfinal.cco' data set in:

Directory: R:/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA
File name: 20190729_2658_CCObaselineWITHcluster.csv

7. Table One preparation

- 10584 participants – both with and without cluster

- 2658 CCO participants – both with and without cluster

Files for the above were saved in:

R:/CloudyWithAChanceOfPain/analyses/Jess/MainCloudybaselinework/NEWDATA/TABLEONE/manuscript
File name:

- 20190730_10584_tableonewithoutcluster.csv
- 20190730_2658_CCO_tableonewithoutcluster.csv
- 20190730_10584_tableoneWITHcluster.csv

For each of the above categories, I retrieved the summary statistics for beliefs (with and without clusters), days in study, number of days of pain data entry per individual and proportion of total days on which pain data entered per individual.

Appendix 1: Explaining the variable names for categorised free text conditions

Note:

- Any variable that starts with "con." has been coded accordingly with the free text provided by the participants, i.e., this coding work was done manually by Jessica on the csv spreadsheet.
- Any variable that starts with "extra." was derived from the variable "cond_open", which is the free text column for chronic pain conditions. Correct medical terms for the conditions were used to replace the original free text (raw data).

No	Variable Name	Description
1	con.ra	Rheumatoid arthritis
2	con.oa	Osteoarthritis
3	con.spa	Ankylosing spondylitis/axial spondyloarthropathy
4	con.gou	Gout/crystal arthropathy
5	con.artr.unsp	Arthritis unspecified
6	con.cwpfm	Chronic widespread pain/fibromyalgia
7	con.mig	Chronic headache (including migraine)
8	con.np	Neuropathic pain
9	con.other	Other condition
10	con.none	No chronic pain condition
11	userid	
12	cond_open	Original free text (raw data)
13	extra.one	Changed free text (other condition - first)
14	extra.two	Changed free text (other condition - second)
15	extra.three	Changed free text (other condition - third)
16	extra.four	Changed free text (other condition - fourth)
17	extra.conds	Changed free text (other condition – fifth or more)
18	unclear.nodiag	Unclear free text/no diagnosis

Appendix 2: Final baseline data set – descriptions for each variable

No.	Column Names	Note
1	weird.string	
2	sex	gender (non-coded)
3	yob	year of birth
4	post.code	first part of the UK postcode
5	site.head	pain site
6	site.face	pain site
7	site.mouth.jaws	pain site
8	site.neck.or.shoulder.pain	pain site
9	site.back.pain	pain site
10	site.stomach.or.abdominal.pain	pain site
11	site.hip.pain	pain site
12	site.knee.pain	pain site
13	site.hands	pain site
14	site.feet	pain site
15	site.multi	pain site
16	site.all	pain site
17	cond.final.none	condition (none)
18	condsum.ra	condition (ra)
19	condsum.oa	condition (oa)
20	condsum.spa	condition (spa)
21	condsum.gou	condition (gout)
22	condsum.artr.unsp	condition (arthritis unspecified)
23	condsum.cwpfm	condition (cwpfm)
24	condsum.mig	condition (migraine)
25	condsum.np	condition (neuropathic pain)
26	condtemp.other	*(column for previous coding)* **NOT FOR USE**
27	drug.final.none	painkiller med (none)
28	drugsum.pcmol	painkiller med (paracetamol)
29	drugsum.nsaids	painkiller med (nsaids)
30	drugsum.simple.anal	painkiller med (simple analgesia)
31	drugsum.weak.opiates	painkiller med (weak opiates)
32	drugsum.strong.opiates	painkiller med (strong opiates)
33	drugsum.for.no	painkiller med (neuropathic pain)
34	drugsum.coded.other	painkiller med (other - coded)
35	drug.other	painkiller med (other - free text)

No.	Column Names	Note
36	extra.drug.other1	new column for each painkiller med (other - free text)
37	extra.drug.other2	new column for each painkiller med (other - free text)
38	extra.drug.other3	new column for each painkiller med (other - free text)
39	extra.drug.other4	new column for each painkiller med (other - free text)
40	extra.drug.other5	new column for each painkiller med (other - free text)
41	drug.dmard.none	dmard (none)
42	drugsum2.ster	dmard (steroids)
43	drugsum2.dmard	dmard (synthetic dmard)
44	drugsum2.bdmard	dmard (biologic dmard)
45	drugsum2.coded.dmard.other	dmard (other - coded)
46	drug.other.dmard	dmard (other - free text)
47	extra.drug.dmard.other1	new column for each dmard (other - free text)
48	extra.drug.dmard.other2	new column for each dmard (other - free text)
49	belief	belief (scale 0-10)
50	belief.rain	belief (rain) - *non-coded*
51	belief.cold	belief (cold) - *non-coded*
52	belief.hot	belief (hot) - *non-coded*
53	belief.dbp	belief (change in pressure) - *non-coded*
54	belief.dt	belief (change in temperature) - *non-coded*
55	belief.other	belief (other) *free text*
56	via	
57	via.other	
58	userid	userid
59	portalid	
60	region	
61	redirect_to	
62	start.date..utc.	
63	submit.date..utc.	
64	network.id	
65	day	
66	Do you have Osteoarthritis?	*(original OA dataset)* **NOT FOR USE**
67	con.ra	*(previously coded condition dataset)* **NOT FOR USE**
68	con.oa	*(previously coded condition dataset)* **NOT FOR USE**
69	con.spa	*(previously coded condition dataset)* **NOT FOR USE**
70	con.gou	*(previously coded condition dataset)* **NOT FOR USE**
71	con.artr.unsp	*(previously coded condition dataset)* **NOT FOR USE**
72	con.cwpfm	*(previously coded condition dataset)* **NOT FOR USE**
73	con.mig	*(previously coded condition dataset)* **NOT FOR USE**

No.	Column Names	Note
74	con.np	*(previously coded condition dataset)* **NOT FOR USE**
75	con.other	*(previously coded condition dataset)* **NOT FOR USE**
76	cond_open	condition(other - free text)
77	extra.one	new column for each condition(other - free text)
78	extra.two	new column for each condition(other - free text)
79	extra.three	new column for each condition(other - free text)
80	extra.four	new column for each condition(other - free text)
81	extra.conds	new column for each condition(other - free text)
82	unclear.nodiag	new column for each condition(other - free text) - unclear/no diagnosis
83	condsum.pain.other	condition (other - coded) *same as cond.final.other*
84	cond.final.other	condition (other - coded) *same as condsum.final.other*
85	totalcond	total for rowSums of columns that contain “condsum.”
86	totaldrugpain	total for rowSums of columns that contain “drugsum.”
87	totaldrugdmard	total for rowSums of columns that contain “drugsum2.”
88	totalpaincond	total for rowSums of pain conditions except condition (other)
89	cond.pain.none	has pain condition (coded 0) or no pain condition (coded 1)

Cloudy with a chance of pain Data dictionary

General objective: Cloudy with a Chance of Pain is the world's first smartphone-based study to investigate the association between weather and chronic pain. Data collection for Cloudy began in January 2016 and ended in April 2017. Five million pieces of symptom data were submitted over 15 months alongside comprehensive weather data from across the UK.

Principal investigator: Prof William Dixon

Variable name	**Variable description**	**Additional information**
UserId	Unique participant identifier	
Baseline Data		
Sex	Question in baseline questionnaire: Are you Male or Female?	0=male, 1= female; there is not a recorded questionnaire entry if the participant did not record a response her
Age	Age derived from year of birth field in baseline questionnaire	Numerical, except: .a = year of birth reported such that participant is >116 years (age of oldest woman alive); probable error .b= year of birth reported such that age is <0; probable error
belief	Question in baseline questionnaire: How likely do you think it is that the weather is associated with pain?	0-10 (10 point scale anchored at 0 “not at all likely” and 10 “extremely likely”)
Site_of_pain*	Question in baseline questionnaire: Where is your pain?	1=Mouth or jaw, 2= Neck or shoulder, 3= Back pain, 4=Stomach or abdominal, 5= Hip pain, 6=Knee pain, 7=Hands, 9=Feet, 10=Pain at multiple site, 11=Pain all over body
Condition_chronicc_pain*	Question in baseline questionnaire: Has your doctor ever told you that you have any of the following conditions?	1= No pain, 2=Rheumatoid arthritis, 3=Osteoarthritis, 4=Spondyloarthropathy, 5= Gout, 6=Unspecific arthritis, 7=Fibromyalgia, 8=Chronic headache, 9=Neuropathic pain, 10= other

Medication*	Question in baseline questionnaire: What types of medication are you currently using?	1=No meds,2=Paracetamol, 3=NSAIDS, 4=analgesic, 5=opiates week, 6=opiates strong, 7=neuro dugs, 8=other pain killer drugs, 9=no DMARD, 10=sDMARD, 11=bDMARDS,12=other_DMARD
Daily recorded data		
Date	Date that the participant records the symptoms	The value is obtained by adding the difference between UTC and local time to the date that the entry is referring to
HMS	Exact Time of the day that symptom entry is recorded	
PainImpct	Has your pain interfered with your activities today?	1=Not at all, 2= a little bit , 3= somewhat, 4=quite a bit, 5=Very much
SleepQuality	How was your sleep quality last night?	1=Very poor, 2= poor , 3= Fair, 4=Good, 5=Very good
timeSpentOuside	How much times have you spent outside today?	1=None of the day, 2= some of the day, 3= Half of the day, 4= Most of the day, 5= All of the day
wakingUpTired	How did you feel when you woke this morning?	1=Not at all tired, 2= a little bit tired , 3= moderately tired, 4=quite a bit tired, 5=Extremely tired
exercise	How long have you exercised today?	1= No exercise, 2= Less than 30 minutes of light activity, 3= 30+ minutes light activity, 4= Less than 30 minute strenuous activity, 5= 30+ minute strenuous activity
Mood	How has your mood been today?	1= Depressed, 2= Felling low, 3= Not very happy, 4= quite happy, 5= very happy
patientWellbeing	How well did you feel today?	1= very unwell, 2= quite unwell, 3= unwell, 4=Well, 5= Very well
PainSeverity	How severe was your pain today?	1= No pain , 2= Mild Pain, 3= Moderate pain, 4= Severe pain, 5= Very severe pain
Fatigue	How severe was your fatigue today?	1=No fatigue, 2= Mild fatigue, 3= moderate fatigue, 4= Severe fatigue, 5= very severe fatigue
morningStiffness	How stiff did you feel on waking this morning?	1= No stiffness, 2= A little Stiff, 3= Moderately stiff, 4= severe stiff, 5= very sever stiff
Weather Data		
Average_Wind_speed	The average wind speed the participant exposed to during the day (00:00:00-23:59:00	Numerical value or blank in the case of missing information. Missing wind speed values may imply 1. The station cannot provide wind gust information

		2. The wind gust is not recorded for that day only 3. We don't have GPS to link to the nearest weather station 4. The participant is far away from any station in UK (>60KM)
Average_Temperature	The average air temperature the participant exposed to during the day	>>
Average_Dewpt	The average amount moisture in the air during the day	>>
Average_pressure	The average Atmospheric pressure the participant exposed to during the day	>>
Average_RHX	The average relative humidity the participant exposed to during the day	>>
Minimum_Wind_speed	The minimum wind speed the participant exposed to during the day (00:00:00-23:59:00	>>
Minimum_Temperature	The minimum air temperature the participant exposed to during the day	>>
Minimum_Dewpt	The minimum amount moisture in the air during the day	>>
Minimum_pressure	The minimum Atmospheric pressure the participant exposed to during the day	>>
Minimum_RHX	The minimum relative humidity the participant	>>

	exposed to during the day	
maximum_Wind_speed	The maximum wind speed the participant exposed to during the day (00:00:00-23:59:00	>>
maximum_Temperature	The maximum air temperature the participant exposed to during the day	>>
maximum_Dewpt	The maximum amount moisture in the air during the day	>>
maximum_pressure	The maximum Atmospheric pressure the participant exposed to during the day	>>
maximum_RHX	The maximum relative humidity the participant exposed to during the day	>>

*multiple answers are possible and new categories will be added

$ For the weather components we will have 7 day lag data as well

Appendix B: Proportions of Each 5-Point Ordinal Scale for Each Pain Symptoms (*Cloudy with a Chance of Pain*)

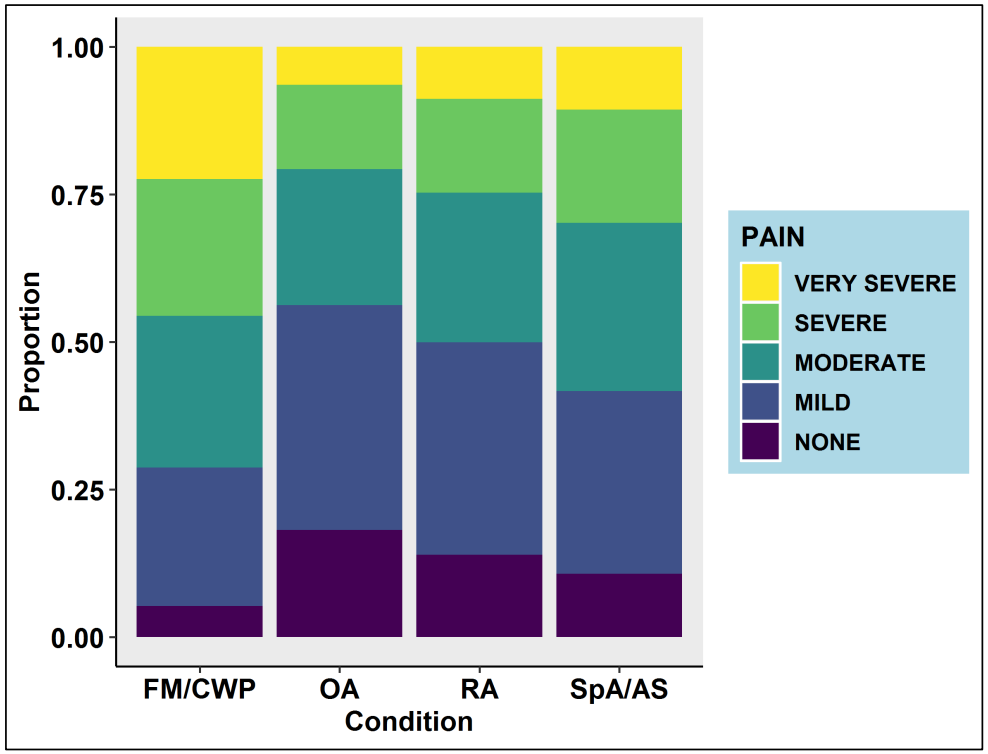
Pain Symptoms

Proportions of each 5-point scale

(A) PAIN SEVERITY

Q: How severe was your pain today?

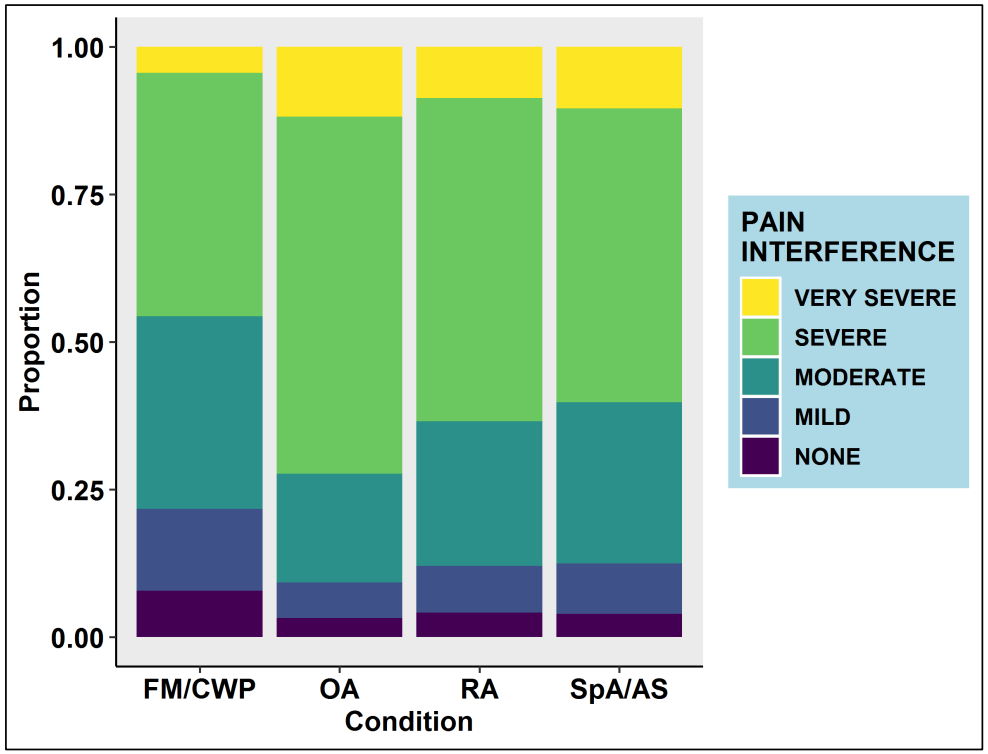
1 – No pain
2 – Mild pain
3 – Moderate pain
4 – Severe pain
5 – Very severe pain



(B) PAIN INTERFERENCE

Q: Has your pain interfered with your activities today?

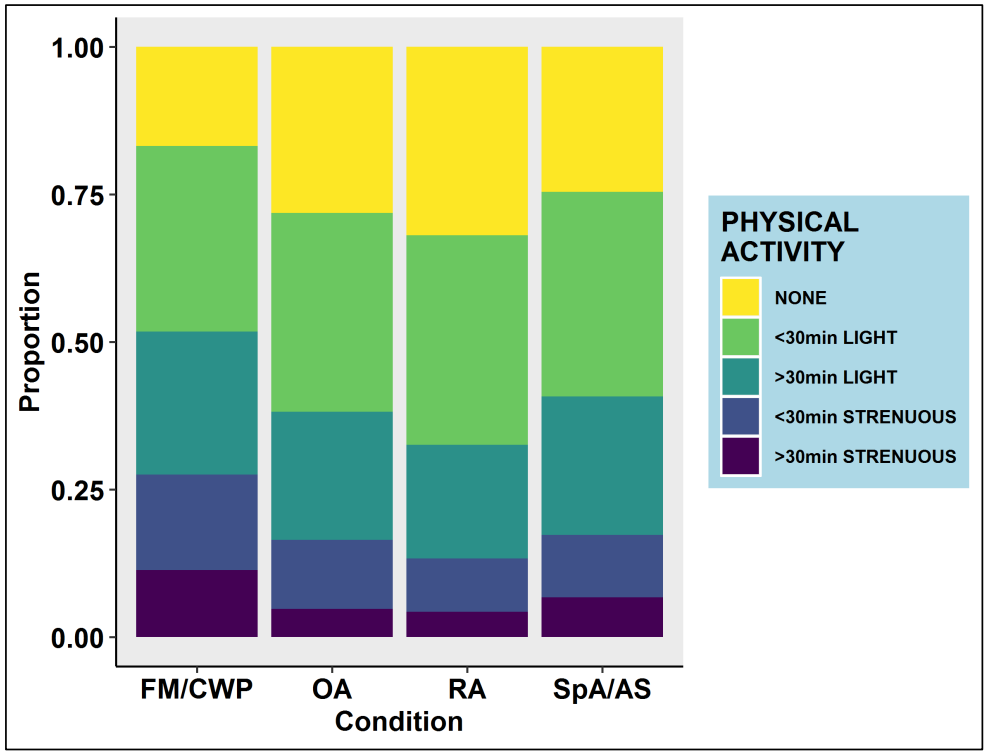
1 – Not at all
2 – A little bit
3 – Somewhat
4 – Quite a bit
5 – Very much



(C) PHYSICAL ACTIVITY

Q: How long have you exercised today?

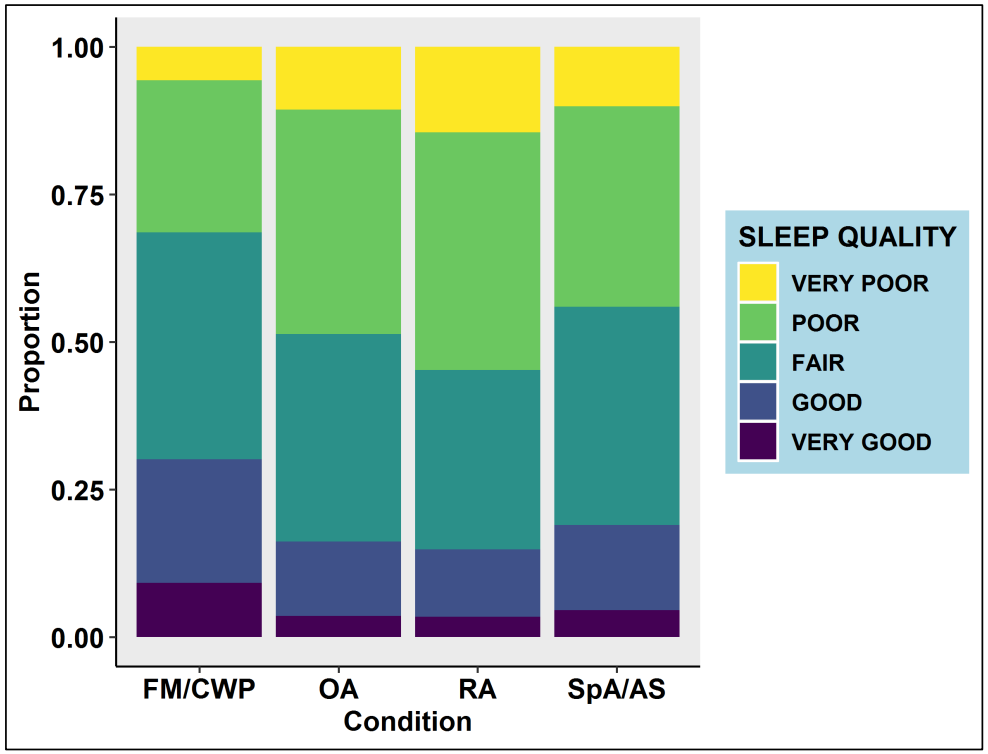
1 – No exercise
2 – Less than 30 minutes of light activity
3 – 30+ minutes of light activity
4 – Less than 30 minutes of strenuous activity
5 – 30+ minutes of strenuous activity



(D) SLEEP QUALITY

Q: How was your sleep quality last night?

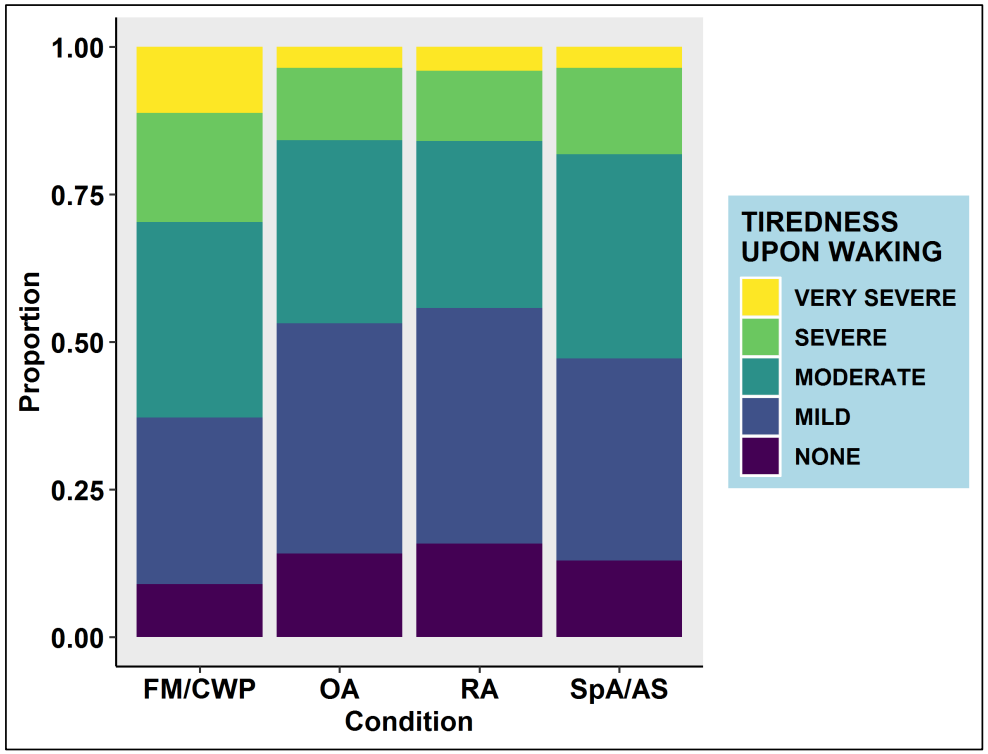
1 – Very poor
2 – Poor
3 – Fair
4 – Good
5 – Very good



(E) TIREDNESS UPON WAKING

Q: How did you feel when you woke this morning?

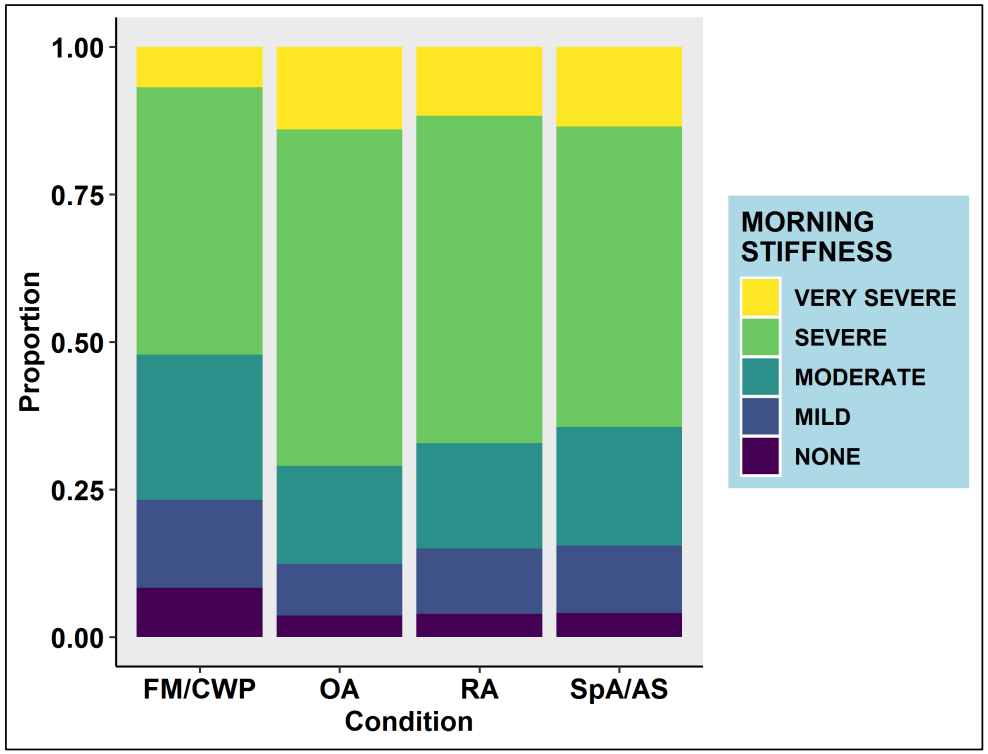
1 – Not at all tired
2 – A little bit tired
3 – Moderately tired
4 – Quite a bit tired
5 – Extremely tired



(F) MORNING STIFFNESS

Q: How stiff did you feel on waking this morning?

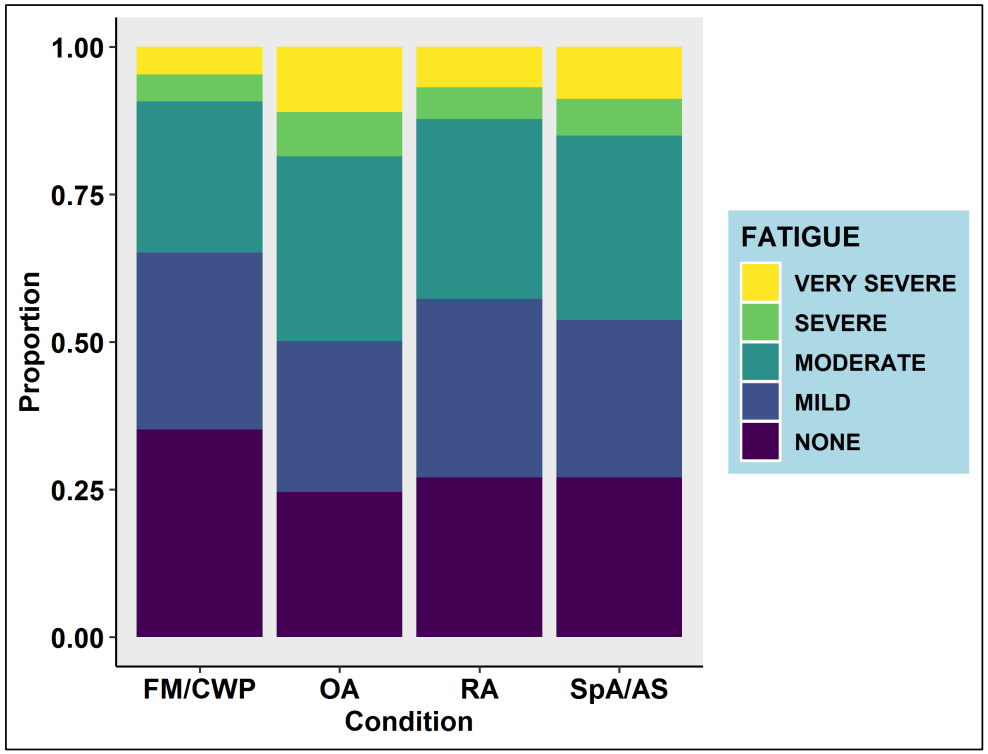
1 – No stiffness
2 – A little stiff
3 – Moderately stiff
4 – Severely stiff
5 – Very severely stiff



(G) FATIGUE

Q: How severe was your fatigue today?

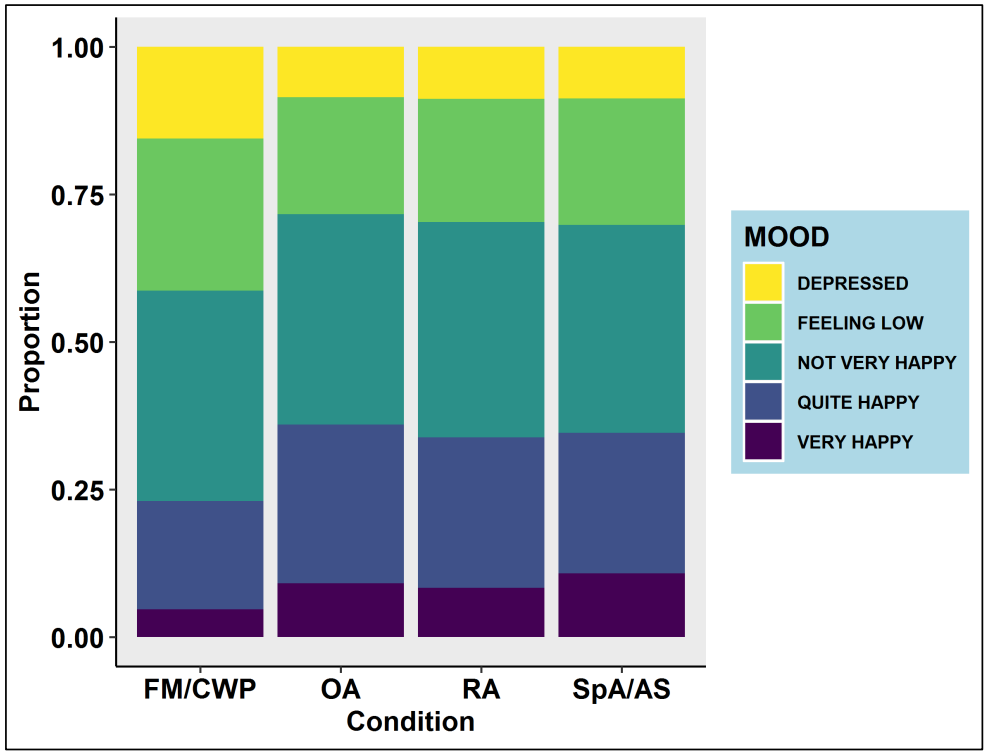
1 – No fatigue
2 – Mild fatigue
3 – Moderate fatigue
4 – Severe fatigue
5 – Very severe fatigue



(H) MOOD

Q: How has your mood been today?

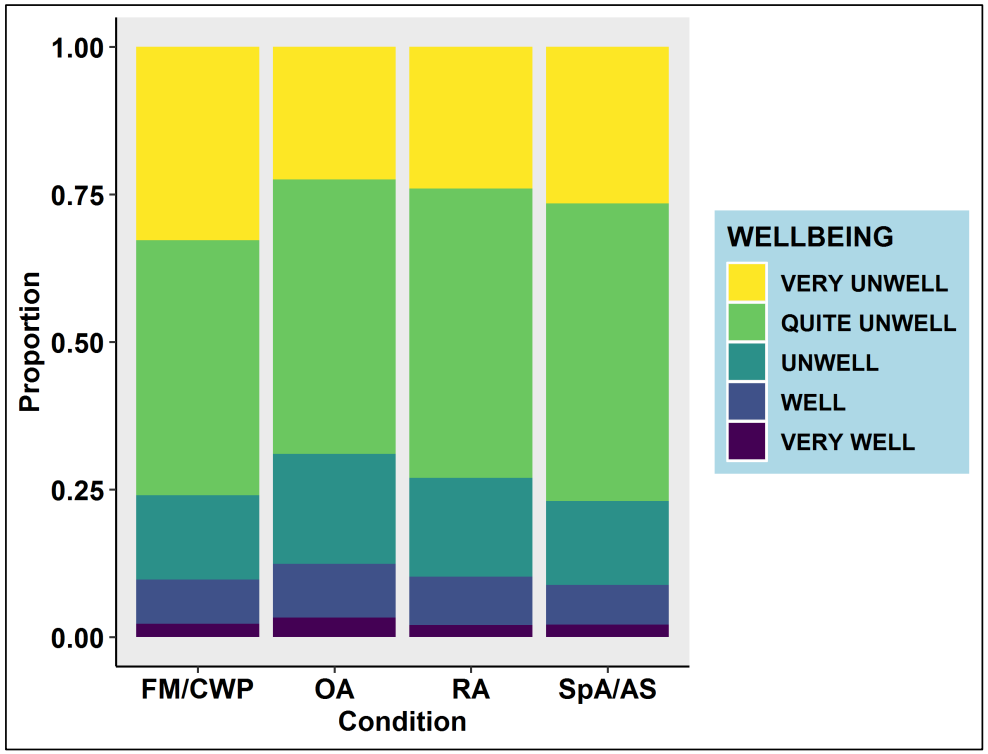
1 – Depressed
2 – Feeling low
3 – Not very happy
4 – Quite happy
5 – Very happy

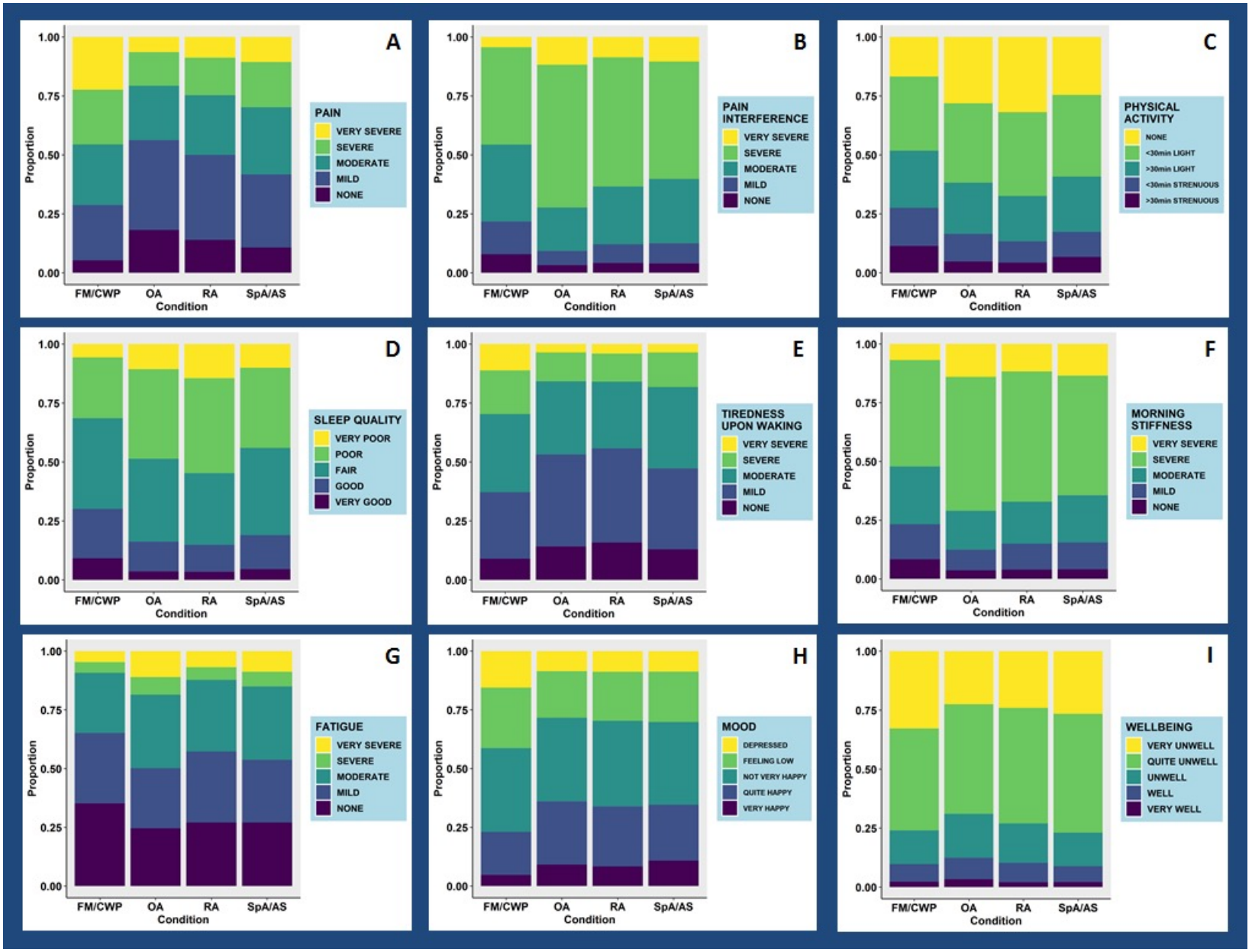


(I) WELLBEING

Q: How well did you feel today?

1 – Very unwell
2 – Quite unwell
3 – Unwell
4 – Well
5 – Very well




A
Proportion
1.00
0.75
0.50
0.25
0.00
FM/CWP
OA
RA
SpA/AS
Condition
PAIN
VERY SEVERE
SEVERE
MODERATE
MILD
NONE
B
PAIN INTERFERENCE
VERY SEVERE
SEVERE
MODERATE
MILD
NONE
C
PHYSICAL ACTIVITY
NONE
<30min LIGHT
>30min LIGHT
<30min STRENUOUS
>30min STRENUOUS
D
SLEEP QUALITY
VERY POOR
POOR
FAIR
GOOD
VERY GOOD
E
TIREDNESS UPON WAKING
VERY SEVERE
SEVERE
MODERATE
MILD
NONE
F
MORNING STIFFNESS
VERY SEVERE
SEVERE
MODERATE
MILD
NONE
G
FATIGUE
VERY SEVERE
SEVERE
MODERATE
MILD
NONE
H
MOOD
DEPRESSED
FEELING LOW
NOT VERY HAPPY
QUITE HAPPY
VERY HAPPY
I
WELLBEING
VERY UNWELL
QUITE UNWELL
UNWELL
WELL
VERY WELL

Appendix C: Statistical Code Scripts for Chapter 4

Single-Disease Pain Analysis

Code

Jessica Pisaniello

24/01/2020

This is R markdown for ARA 2020 abstract.

Original R script from R markdown file [20200106_RAvsOA_pain copy.Rmd].

```
#### R LIBRARY
library(knitr)
library(kableExtra)
library(tidyverse)
library(psycho)
library(gridExtra)
library(data.table)
library(tableone)
library(viridisLite)
library(viridis)
library(lattice)
library(lme4)
library(nlme)
library(sjPlot)
library(sjmisc)
library(effects)
library(ggeffects)
library(ggthemes)
library(lcmm)
library(ggpubr)
library(seqHMM)
library(TraMineR)
library(reshape2)
library(DataCombine)
```

```
#### ORIGINAL BASELINE DATA SET

# Step 1: Load baseline data
#setwd("R:/Projects/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA") ## UoM remote working directory
setwd("~/Cloudy") ##Incline working directory
bsl <-  fread("210519_final_correctedbaseline_11822.csv", header = T) ##The overall baseline data for 11822 participants

## Step 1a: Subsetting above for PhD baseline data
baseline <- bsl %>%
  dplyr::select(2:25,84,26:48,58,65,76:82)
#colnames(baseline) ##To describe the variables that I have selected (from the above) in this new baseline dataset
```

```
#### ORIGINAL MOTIF DATA SET - DATA CLEANING

## Step 2: Load motif data set
#setwd("R:/Projects/CloudyWithAChanceOfPain/common_files/data/external/FINALDATA") ##
UoM remote working directory
setwd("~/Cloudy") ##Incline working directory
motif <- fread("20190719_cloudy-data_motif-weather-bsl_794007-rows_10584_userids.csv"
, header = T) ##The overall motif data for 10584 participants

## Step 2a: Subsetting above for PhD motif data
cleanmotif <- motif %>%
  dplyr::select(1:2, 4:13)
#colnames(cleanmotif) ##To describe the variables that I have selected (from the abov
e) in this new motif dataset

## Step 2b: Cleaning the motif data
cleanmotif[is.na(cleanmotif)]<-0 ##I have changed NA to 0 to facilitate the row sum f
or below
cleanmotif$ind<-rowSums(cleanmotif[,3:12]) ##Row sum for all of the 10 motif items

cleanmotif <- subset(cleanmotif, cleanmotif$ind!=0) %>%
  dplyr::select(-ind) ##I have removed observations with no motif data at all
setDT(cleanmotif) #466325 obs (with motif data)

## Step 2c: To calculate days in study and number of entry
cleanmotif<-cleanmotif[order(UserId, day)]

cleanmotif<-cleanmotif %>%
  group_by(UserId) %>%
  dplyr::mutate(YMD=as.Date(day,format="%Y-%m-%d"),
                first = dplyr::first(YMD),
                last = dplyr::last(YMD))

cleanmotif<-cleanmotif%>%
  mutate(daysinstudy = difftime(last,first,units="days")+1)

cleanmotif<-cleanmotif%>%
  group_by(UserId) %>%
  dplyr::mutate(
    nentry = dplyr::n_distinct(YMD))
setDT(cleanmotif)

## Step 2d: Change calander time to process time to obtain time (to set time of first
entry and onwards)
cleanmotif <- cleanmotif %>%
  group_by(UserId) %>%
  mutate(Date=as.Date(day,format="%Y-%m-%d"),
         prev.entry_date = c(0, diff(Date)))
cleanmotif <- ungroup(cleanmotif)

cleanmotif<-cleanmotif %>%
  group_by(UserId) %>%
  mutate(time = cumsum(prev.entry_date))
setDT(cleanmotif)

## Step 2e: Change colnames for UserId to userid to facilitate left join for data mer
ging later
```

```
colnames(cleanmotif)[colnames(cleanmotif)=="UserId"] <- "userid"

### EXTRA: To confirm number of participant in my cleanmotif is 10584
#cleanmotif %>%
  #summarise(total = n_distinct(userid)) ## 10584 participants
```

```
#### BASELINE DATA SET FOR ANALYSIS - DATA CLEANING

## Step 3: Get baseline data for 10584 participants
cloudybsl <- baseline %>%
  filter(baseline$userid %in% cleanmotif$userid)

## Step 3a: Change column names for conditions other and none
colnames(cloudybsl)[colnames(cloudybsl)=="cond.final.none"]<-"condsum.final.none"
colnames(cloudybsl)[colnames(cloudybsl)=="cond.final.other"]<-"condsum.final.other"

## Step 3b: Create a new column for multicond
cloudybsl$multicond<-cloudybsl %>%
  dplyr::select(contains("condsum.")) %>%
  rowSums(na.rm=T) ##This facilitates analysis for single condition vs multiple condi
tions

## Step 3c: Change gender female = 1, male = 0
cloudybsl$sex <- ifelse(cloudybsl$sex=="Female", 1, 0)

## Step 3d: Calculate the age
cloudybsl$yob[!cloudybsl$yob %in% c(1900:2000)] <- NA

#sum(is.na(cloudybsl$yob)) #401 without proper yob entered by participants

cloudybsl$yrentry <- year(as.Date(cloudybsl$day, format = "%d/%m/%Y"))
cloudybsl$age <- as.numeric(cloudybsl$yrentry) - as.numeric(cloudybsl$yob) ##To obtai
n the age
```

```
#### BASELINE DATA SET FOR RHEUMATOID ARTHRITIS (RA), OSTEOARTHRITIS (OA) AND FIBROMY
ALGIA (CWPFM) AND SPONDYLOARTHROPATHY (SPA)

## Step 4: Main baseline data for single condition
singlecondbsl<-cloudybsl %>%
  filter(multicond==1) #5976 participants

## Step 4b: Get baseline data for each condition of interest
###### RA only cohort ######
ra.bsl<-singlecondbsl %>%
  filter(condsum.ra==1) #921 participants

###### OA only cohort ######
oa.bsl<-singlecondbsl %>%
  filter(condsum.oa==1) #758 participants

###### CWPFM only cohort ######
cwpfm.bsl<-singlecondbsl %>%
  filter(condsum.cwpfm==1) #630 participants

###### SpA only cohort ######
spa.bsl<-singlecondbsl %>%
  filter(condsum.spa==1) #216 participants
```

```
#### CLEAN MOTIF DATA SET FOR ANALYSIS

## Step 4c: Get motif data for the above for a one month period for participants who
 remained in the study for 30 days or more (to exclude tourists)
##### RA #####
ra.motif <- cleanmotif %>%
  filter(cleanmotif$userid %in% ra.bsl$userid) #42840 obs (overall motif data for 921
participants)
#ra.motif %>%
  #summarise(total = n_distinct(userid)) #921 participants

ra.motif <- ra.motif %>%
  filter(daysinstudy>29) #40321 obs (participants who remained in the study for 30 da
ys or more)
#ra.motif %>%
  #summarise(total = n_distinct(userid)) #425 participants
#summary(ra.motif$nentry)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#2.0    96.0   172.0   194.9   291.0   429.0


ra.finalmotif <- ra.motif %>%
  filter(time<31) #8752 obs (above criteria + first one month motif data)
#ra.finalmotif %>%
  #summarise(total = n_distinct(userid)) #425 participants
#summary(ra.finalmotif$nentry)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#2      39      89     118     169     429


##### OA #####
oa.motif <- cleanmotif %>%
  filter(cleanmotif$userid %in% oa.bsl$userid) #45606 obs (overall motif data for 758
participants)
#oa.motif %>%
  #summarise(total = n_distinct(userid)) #758 participants

oa.motif <- oa.motif %>%
  filter(daysinstudy>29) #43867 obs (participants who remained in the study for 30 da
ys or more)
#oa.motif %>%
  #summarise(total = n_distinct(userid)) #409 participants
#summary(oa.motif$nentry)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#2.0   104.0   173.0   202.4   299.0   453.0

oa.finalmotif <- oa.motif %>%
  filter(time<31) #8775 obs (above criteria + first one month motif data)
#oa.finalmotif %>%
  #summarise(total = n_distinct(userid)) #409 participants
#summary(oa.finalmotif$nentry)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#2.0    48.0    94.0   127.1   170.0   453.0


##### CWPFM #####
cwpfm.motif <- cleanmotif %>%
```

```
  filter(cleanmotif$userid %in% cwpfm.bsl$userid) #21544 obs (overall motif data for
 630 participants)
#cwpfm.motif %>%
  #summarise(total = n_distinct(userid)) #630 participants

cwpfm.motif <- cwpfm.motif %>%
  filter(daysinstudy>29) #19708 obs (participants who remained in the study for 30 da
ys or more)
#cwpfm.motif %>%
  #summarise(total = n_distinct(userid)) #255 participants
#summary(cwpfm.motif$nentry)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#2.0    73.0   165.0   184.7   264.0   436.0

cwpfm.finalmotif <- cwpfm.motif %>%
  filter(time<31) #4992 obs (above criteria + first one month motif data)
#cwpfm.finalmotif %>%
  #summarise(total = n_distinct(userid)) #255 particpants
#summary(cwpfm.finalmotif$nentry)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#2.00   32.00   60.00   99.19  146.00  436.00

##### SpA #####
spa.motif <- cleanmotif %>%
  filter(cleanmotif$userid %in% spa.bsl$userid) #9429 obs (overall motif data for 216
participants)
#spa.motif %>%
  #summarise(total = n_distinct(userid)) #216 participants

spa.motif <- spa.motif %>%
  filter(daysinstudy>29) #8901 obs (participants who remained in the study for 30 day
s or more)
#spa.motif %>%
  #summarise(total = n_distinct(userid)) #100 participants
#summary(spa.motif$nentry)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#4.0   109.0   179.0   187.8   268.0   402.0

spa.finalmotif <- spa.motif %>%
  filter(time<31) #1979 obs (above criteria + first one month motif data)
#spa.finalmotif %>%
  #summarise(total = n_distinct(userid)) #100 particpants
#summary(spa.finalmotif$nentry)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#4.0    34.0    87.0   114.1   175.0   402.0

singledisease.first30days <- rbind(ra.motif, spa.motif, oa.motif, cwpfm.motif)
#singledisease.first30days %>%
  #summarise(total=n_distinct(userid)) #1189 participants
singledisease.first30days <- singledisease.first30days %>%
  group_by(userid) %>%
  mutate(nentry2 = n_distinct(YMD))
#summary(singledisease.first30days$nentry2)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#2.0    94.0   171.0   195.5   293.0   453.0

singledisease.first1mth <- rbind(ra.finalmotif, spa.finalmotif, oa.finalmotif, cwpfm.
finalmotif)
```

```
#singledisease.first1mth %>%
  #summarise(total = n_distinct(userid)) #1189 participants
singledisease.first1mth <- singledisease.first1mth %>%
  group_by(userid) %>%
  mutate(nentry2 = n_distinct(YMD))
#summary(singledisease.first1mth$nentry2)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
#1.00   21.00   26.00   24.29   30.00   31.00
```

BASELINE CHARACTERISTICS

RA

Code Output

	Overall
n	425
sex = 1 (%)	354 (83.3)
age (mean (SD))	49.17 (11.70)
site.head = 1 (%)	21 (4.9)
site.face = 1 (%)	6 (1.4)
site.mouth.jaws = 1 (%)	59 (13.9)
site.neck.or.shoulder.pain = 1 (%)	231 (54.4)
site.back.pain = 1 (%)	122 (28.7)
site.stomach.or.abdominal.pain = 1 (%)	20 (4.7)
site.hip.pain = 1 (%)	175 (41.2)
site.knee.pain = 1 (%)	251 (59.1)
site.hands = 1 (%)	324 (76.2)
site.feet = 1 (%)	267 (62.8)
site.multi = 1 (%)	197 (46.4)
site.all = 1 (%)	45 (10.6)
drug.final.none = 1 (%)	38 (8.9)
drugsum.pcmol = 1 (%)	182 (42.8)
drugsum.nsaids = 1 (%)	270 (63.5)

	Overall
drugsum.simple.anal = 1 (%)	124 (29.2)
drugsum.weak.opiates = 1 (%)	93 (21.9)
drugsum.strong.opiates = 1 (%)	30 (7.1)
drugsum.for.no = 1 (%)	22 (5.2)
drugsum.coded.other = 1 (%)	10 (2.4)
drug.dmard.none = 1 (%)	59 (13.9)
drugsum2.ster = 1 (%)	80 (18.8)
drugsum2.dmard = 1 (%)	328 (77.2)
drugsum2.bdmard = 1 (%)	110 (25.9)
drugsum2.coded.dmard.other = 0 (%)	425 (100.0)

SpA

Code Output

	Overall
n	100
sex = 1 (%)	65 (65.0)
age (mean (SD))	44.90 (11.96)
site.head = 1 (%)	9 (9.0)
site.face = 1 (%)	2 (2.0)
site.mouth.jaws = 1 (%)	13 (13.0)
site.neck.or.shoulder.pain = 1 (%)	64 (64.0)
site.back.pain = 1 (%)	88 (88.0)
site.stomach.or.abdominal.pain = 1 (%)	11 (11.0)
site.hip.pain = 1 (%)	57 (57.0)
site.knee.pain = 1 (%)	45 (45.0)
site.hands = 1 (%)	40 (40.0)

	Overall
site.feet = 1 (%)	35 (35.0)
site.multi = 1 (%)	41 (41.0)
site.all = 1 (%)	5 (5.0)
drug.final.none = 1 (%)	7 (7.0)
drugsum.pcmol = 1 (%)	38 (38.0)
drugsum.nsaids = 1 (%)	75 (75.0)
drugsum.simple.anal = 1 (%)	22 (22.0)
drugsum.weak.opiates = 1 (%)	25 (25.0)
drugsum.strong.opiates = 1 (%)	13 (13.0)
drugsum.for.no = 1 (%)	13 (13.0)
drugsum.coded.other = 1 (%)	1 (1.0)
drug.dmard.none = 1 (%)	50 (50.0)
drugsum2.ster = 1 (%)	5 (5.0)
drugsum2.dmard = 1 (%)	25 (25.0)
drugsum2.bdmard = 1 (%)	28 (28.0)
drugsum2.coded.dmard.other = 0 (%)	100 (100.0)

OA

Code Output

	Overall
n	409
sex = 1 (%)	340 (83.1)
age (mean (SD))	57.01 (10.65)
site.head = 1 (%)	13 (3.2)
site.face = 1 (%)	4 (1.0)
site.mouth.jaws = 1 (%)	14 (3.4)

	Overall
site.neck.or.shoulder.pain = 1 (%)	151 (36.9)
site.back.pain = 1 (%)	186 (45.5)
site.stomach.or.abdominal.pain = 1 (%)	15 (3.7)
site.hip.pain = 1 (%)	207 (50.6)
site.knee.pain = 1 (%)	282 (68.9)
site.hands = 1 (%)	215 (52.6)
site.feet = 1 (%)	147 (35.9)
site.multi = 1 (%)	67 (16.4)
site.all = 1 (%)	13 (3.2)
drug.final.none = 1 (%)	35 (8.6)
drugsum.pcmol = 1 (%)	220 (53.8)
drugsum.nsaids = 1 (%)	261 (63.8)
drugsum.simple.anal = 1 (%)	114 (27.9)
drugsum.weak.opiates = 1 (%)	101 (24.7)
drugsum.strong.opiates = 1 (%)	26 (6.4)
drugsum.for.no = 1 (%)	31 (7.6)
drugsum.coded.other = 1 (%)	25 (6.1)
drug.dmard.none = 1 (%)	369 (90.2)
drugsum2.ster = 1 (%)	15 (3.7)
drugsum2.dmard = 1 (%)	16 (3.9)
drugsum2.bdmard = 1 (%)	8 (2.0)
drugsum2.coded.dmard.other = 1 (%)	2 (0.5)

CWPFM

Code Output

	Overall

	Overall
n	255
sex = 1 (%)	228 (89.4)
age (mean (SD))	41.44 (10.89)
site.head = 1 (%)	70 (27.5)
site.face = 1 (%)	37 (14.5)
site.mouth.jaws = 1 (%)	71 (27.8)
site.neck.or.shoulder.pain = 1 (%)	157 (61.6)
site.back.pain = 1 (%)	165 (64.7)
site.stomach.or.abdominal.pain = 1 (%)	64 (25.1)
site.hip.pain = 1 (%)	150 (58.8)
site.knee.pain = 1 (%)	149 (58.4)
site.hands = 1 (%)	132 (51.8)
site.feet = 1 (%)	112 (43.9)
site.multi = 1 (%)	169 (66.3)
site.all = 1 (%)	121 (47.5)
drug.final.none = 1 (%)	14 (5.5)
drugsum.pcmol = 1 (%)	118 (46.3)
drugsum.nsaids = 1 (%)	133 (52.2)
drugsum.simple.anal = 1 (%)	86 (33.7)
drugsum.weak.opiates = 1 (%)	107 (42.0)
drugsum.strong.opiates = 1 (%)	29 (11.4)
drugsum.for.no = 1 (%)	108 (42.4)
drugsum.coded.other = 1 (%)	47 (18.4)
drug.dmard.none = 1 (%)	233 (91.4)
drugsum2.ster = 1 (%)	5 (2.0)
drugsum2.dmard = 1 (%)	12 (4.7)
drugsum2.bdmard = 1 (%)	5 (2.0)

	Overall
drugsum2.coded.dmard.other = 1 (%)	1 (0.4)

OVERALL BASELINE CHARACTERISTICS

Code Output

	Overall
n	1189
sex = 1 (%)	987 (83.0)
age (mean (SD))	49.89 (12.66)
site.head = 1 (%)	113 (9.5)
site.face = 1 (%)	49 (4.1)
site.mouth.jaws = 1 (%)	157 (13.2)
site.neck.or.shoulder.pain = 1 (%)	603 (50.7)
site.back.pain = 1 (%)	561 (47.2)
site.stomach.or.abdominal.pain = 1 (%)	110 (9.3)
site.hip.pain = 1 (%)	589 (49.5)
site.knee.pain = 1 (%)	727 (61.1)
site.hands = 1 (%)	711 (59.8)
site.feet = 1 (%)	561 (47.2)
site.multi = 1 (%)	474 (39.9)
site.all = 1 (%)	184 (15.5)
drug.final.none = 1 (%)	94 (7.9)
drugsum.pcmol = 1 (%)	558 (46.9)
drugsum.nsaids = 1 (%)	739 (62.2)
drugsum.simple.anal = 1 (%)	346 (29.1)
drugsum.weak.opiates = 1 (%)	326 (27.4)
drugsum.strong.opiates = 1 (%)	98 (8.2)

	Overall
drugsum.for.no = 1 (%)	174 (14.6)
drugsum.coded.other = 1 (%)	83 (7.0)
drug.dmard.none = 1 (%)	711 (59.8)
drugsum2.ster = 1 (%)	105 (8.8)
drugsum2.dmard = 1 (%)	381 (32.0)
drugsum2.bdmard = 1 (%)	151 (12.7)
drugsum2.coded.dmard.other = 1 (%)	3 (0.3)

DESCRIPTIVE SUMMARY OF PAIN SEVERITY FOR THE FIRST ONE MONTH

SUMMARY FOR RA

Code Output

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.354   2.438   2.514   2.497   2.560   2.655
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.8836  0.9470  0.9739  0.9766  1.0130  1.0520
```

time	N	mean	sd	sem
0	361	2.631579	0.9516068	0.0500846
1	327	2.559633	0.9978307	0.0551802
2	300	2.560000	0.9430613	0.0544477
3	290	2.593103	1.0188300	0.0598278
4	296	2.547297	1.0073254	0.0585496
5	283	2.568905	0.9918218	0.0589577
6	296	2.574324	1.0323002	0.0600012
7	284	2.580986	0.9967036	0.0591435
8	258	2.655039	1.0293333	0.0640835

time	N	mean	sd	sem
9	262	2.461832	0.9571636	0.0591338
10	278	2.514388	1.0043991	0.0602399
11	268	2.608209	1.0347193	0.0632055
12	270	2.514815	1.0301058	0.0626902
13	274	2.394161	0.9710661	0.0586643
14	266	2.387218	0.9216813	0.0565119
15	262	2.438931	0.9357172	0.0578088
16	270	2.462963	0.9738774	0.0592683
17	273	2.498168	0.9553431	0.0578200
18	255	2.525490	1.0186921	0.0637930
19	257	2.354086	0.9494616	0.0592258
20	256	2.437500	0.9801961	0.0612623
21	258	2.515504	1.0520278	0.0654964
22	243	2.522634	1.0217177	0.0655432
23	253	2.454546	0.9568618	0.0601574
24	263	2.395437	0.9746296	0.0600982
25	250	2.440000	0.9391101	0.0593945
26	248	2.383065	0.9234000	0.0586360
27	252	2.436508	0.8836364	0.0556639
28	271	2.516605	0.9263134	0.0562695
29	227	2.440529	0.9499584	0.0630510
30	239	2.422594	0.9445064	0.0610950

SUMMARY FOR SpA

Code Output

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.600   2.669   2.724   2.734   2.798   2.929
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.8044  0.9235  0.9775  0.9834  1.0801  1.1076
```

time	N	mean	sd	sem
0	91	2.736264	1.0092369	0.1057968
1	76	2.750000	0.9398581	0.1078091
2	73	2.835616	0.9720211	0.1137665
3	77	2.688312	1.0164404	0.1158341
4	67	2.850746	1.0766634	0.1315354
5	66	2.833333	0.9540736	0.1174383
6	70	2.928571	1.1075685	0.1323798
7	67	2.671642	1.0500802	0.1282877
8	63	2.777778	1.0842635	0.1366044
9	57	2.859649	1.0426374	0.1381007
10	57	2.719298	0.9775039	0.1294735
11	59	2.762712	0.8971023	0.1167928
12	60	2.650000	0.9356408	0.1207907
13	57	2.666667	0.8309490	0.1100619
14	65	2.830769	0.8762354	0.1086836
15	61	2.622951	0.9515590	0.1218346
16	55	2.600000	0.8520129	0.1148854
17	61	2.721312	0.9332943	0.1194961
18	58	2.724138	0.8744543	0.1148215
19	57	2.719298	0.9956044	0.1318710
20	58	2.689655	0.9770506	0.1282930
21	55	2.727273	0.8488455	0.1144583
22	55	2.818182	1.0902074	0.1470035
23	58	2.724138	0.9136989	0.1199745
24	51	2.607843	1.0968761	0.1535934

time	N	mean	sd	sem
25	50	2.640000	1.0834563	0.1532239
26	59	2.677966	1.1054829	0.1439216
27	59	2.661017	0.9932560	0.1293109
28	59	2.644068	0.8043569	0.1047184
29	55	2.763636	1.0880434	0.1467117
30	49	2.836735	1.1057980	0.1579711

SUMMARY FOR OA

Code Output

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.498   2.560   2.592   2.601   2.618   2.765
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.8903  0.9270  0.9470  0.9464  0.9588  1.0073
```

time	N	mean	sd	sem
0	357	2.764706	1.0031345	0.0530915
1	317	2.757098	0.9617635	0.0540180
2	313	2.616613	0.9336886	0.0527752
3	303	2.577558	0.9312944	0.0535015
4	297	2.545454	0.9857710	0.0572002
5	281	2.530249	0.9141602	0.0545342
6	268	2.641791	0.9707558	0.0592983
7	289	2.560554	0.9559578	0.0562328
8	270	2.637037	0.9571287	0.0582490
9	273	2.571429	0.9487940	0.0574236
10	272	2.584559	0.9450886	0.0573044
11	270	2.688889	0.9978912	0.0607297
12	261	2.551724	0.9126288	0.0564903

time	N	mean	sd	sem
13	272	2.544118	0.9040103	0.0548137
14	273	2.615385	0.9521944	0.0576294
15	264	2.613636	0.9158322	0.0563656
16	250	2.592000	0.9577081	0.0605708
17	258	2.620155	1.0073180	0.0627129
18	260	2.596154	0.9598024	0.0595244
19	250	2.624000	0.9707696	0.0613969
20	247	2.615385	0.9466705	0.0602352
21	263	2.699620	0.9395264	0.0579337
22	257	2.610895	0.9542043	0.0595216
23	261	2.521073	0.9469264	0.0586133
24	250	2.604000	0.9480396	0.0599593
25	247	2.558705	0.9346509	0.0594704
26	241	2.535270	0.9081809	0.0585011
27	249	2.586345	0.9469679	0.0600117
28	255	2.498039	0.9130485	0.0571773
29	261	2.567050	0.8902822	0.0551071
30	254	2.590551	0.9227415	0.0578980

SUMMARY FOR CWPFM

Code Output

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.848   3.000   3.022   3.041   3.074   3.245
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.9307  0.9956  1.0286  1.0320  1.0721  1.1328
```

time	N	mean	sd	sem
0	220	3.245455	0.9478297	0.0639027

time	N	mean	sd	sem
1	185	3.118919	0.9707230	0.0713690
2	193	3.088083	1.0692253	0.0769645
3	181	3.022099	0.9306855	0.0691773
4	166	3.156626	0.9658650	0.0749657
5	171	3.122807	1.0911975	0.0834460
6	164	3.006098	1.0537514	0.0822842
7	159	3.176101	1.0586270	0.0839546
8	168	3.059524	0.9952124	0.0767823
9	160	3.056250	1.0535766	0.0832925
10	158	3.120253	1.1136535	0.0885975
11	154	2.980520	1.0127990	0.0816137
12	149	2.959732	0.9647797	0.0790378
13	156	3.000000	1.0286226	0.0823557
14	151	2.993377	1.1105355	0.0903741
15	151	2.933775	1.0749814	0.0874807
16	146	3.047945	1.0912327	0.0903111
17	149	3.040268	1.0772788	0.0882541
18	145	2.848276	0.9741879	0.0809019
19	139	3.021583	1.0664014	0.0904510
20	143	3.020979	0.9962502	0.0833106
21	131	3.221374	1.0174635	0.0888962
22	140	3.057143	1.0913092	0.0922325
23	142	2.950704	1.0405090	0.0873176
24	137	2.854015	0.9666509	0.0825866
25	137	3.021898	0.9960743	0.0851004
26	142	3.021127	1.1328011	0.0950626
27	138	3.014493	1.0107850	0.0860437

time	N	mean	sd	sem
28	147	3.047619	1.0490265	0.0865222
29	132	3.000000	1.0263699	0.0893340
30	138	3.050725	1.0132153	0.0862506

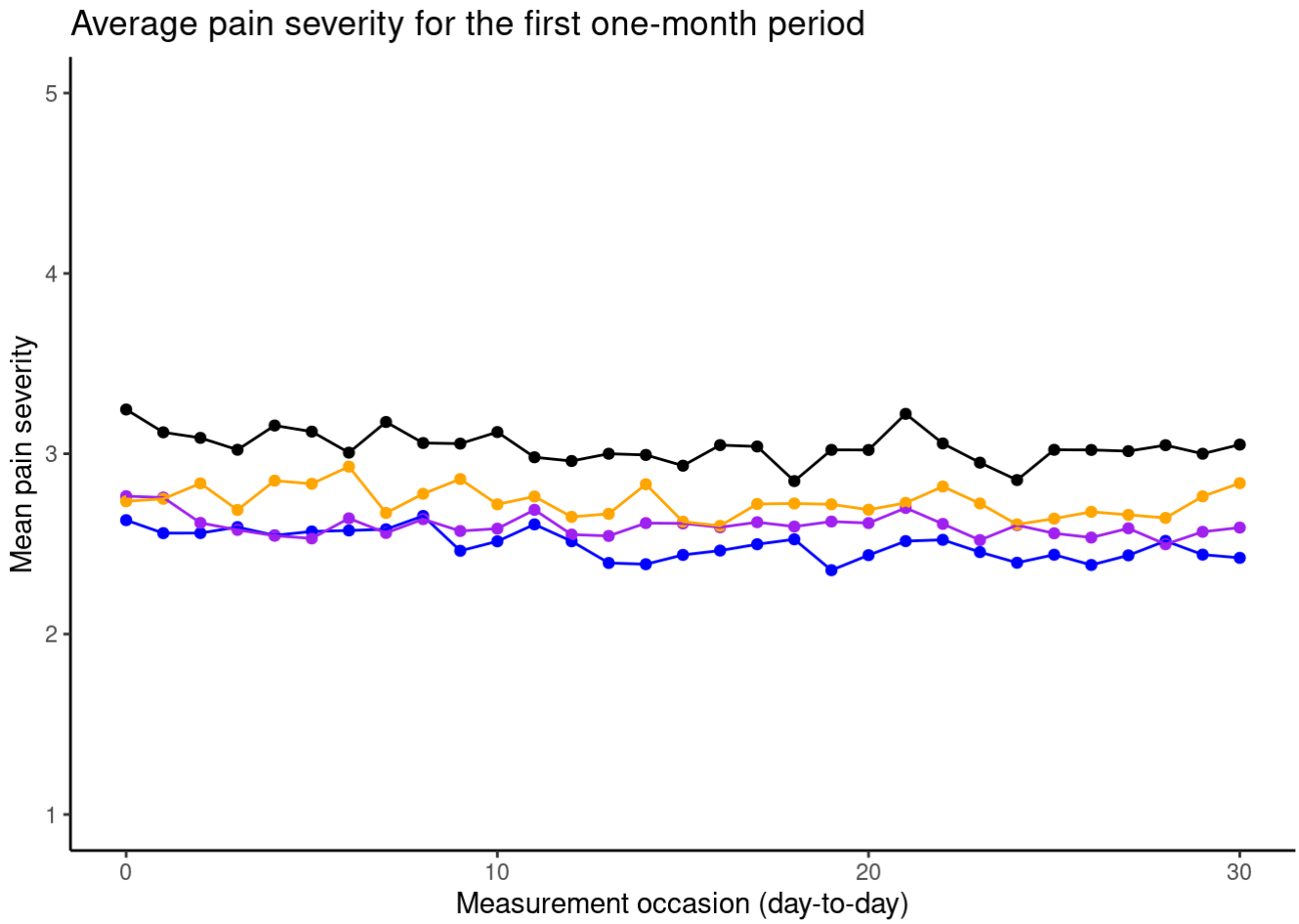
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   21.00   26.00   24.29   30.00   31.00
```

AVERAGE PAIN SEVERITY FOR THE FIRST ONE MONTH

OVERVIEW

- BLUE - RA
- PURPLE - OA
- BLACK - CWPFM
- ORANGE - SpA

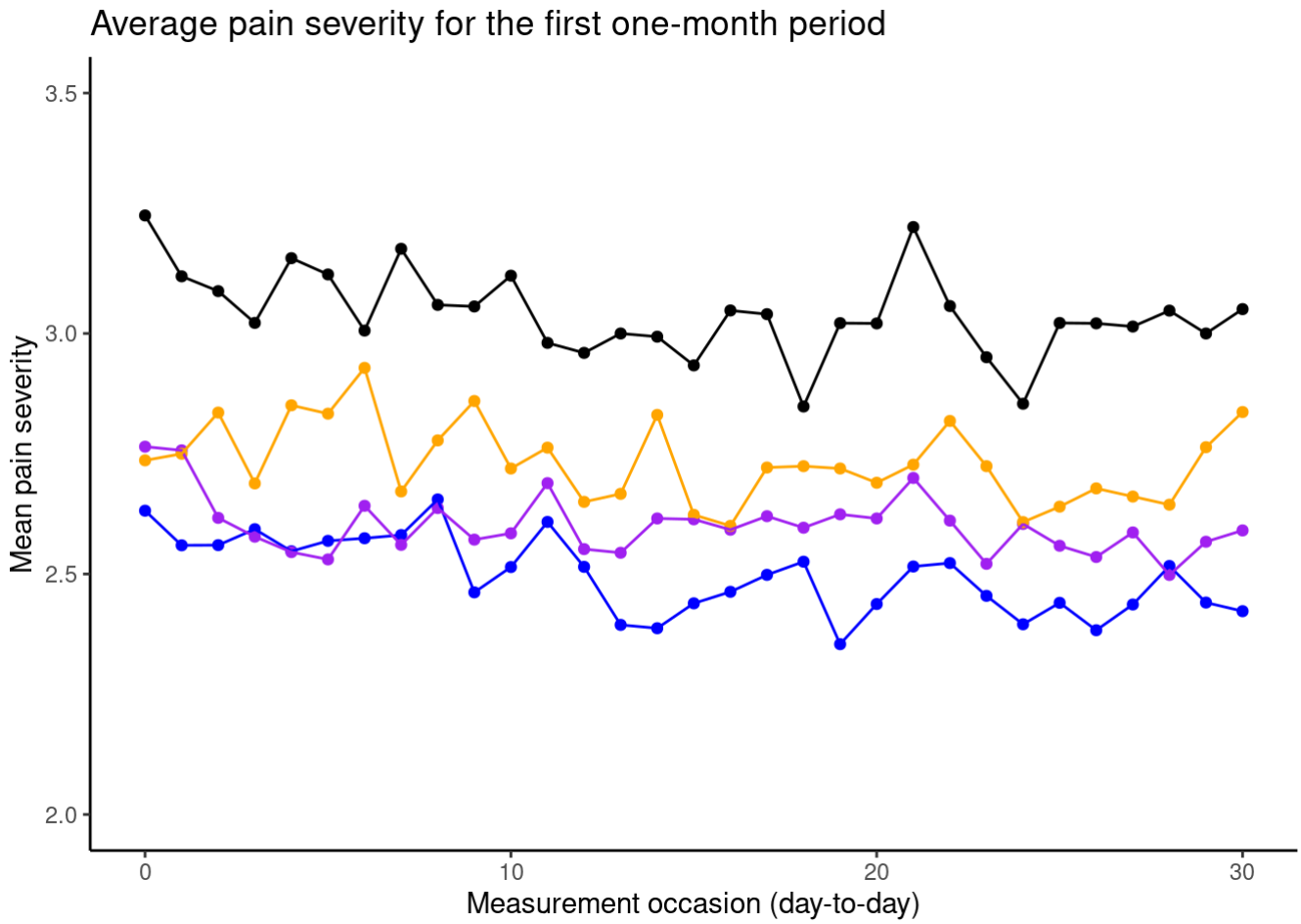
Code Output



CLOSE-UP

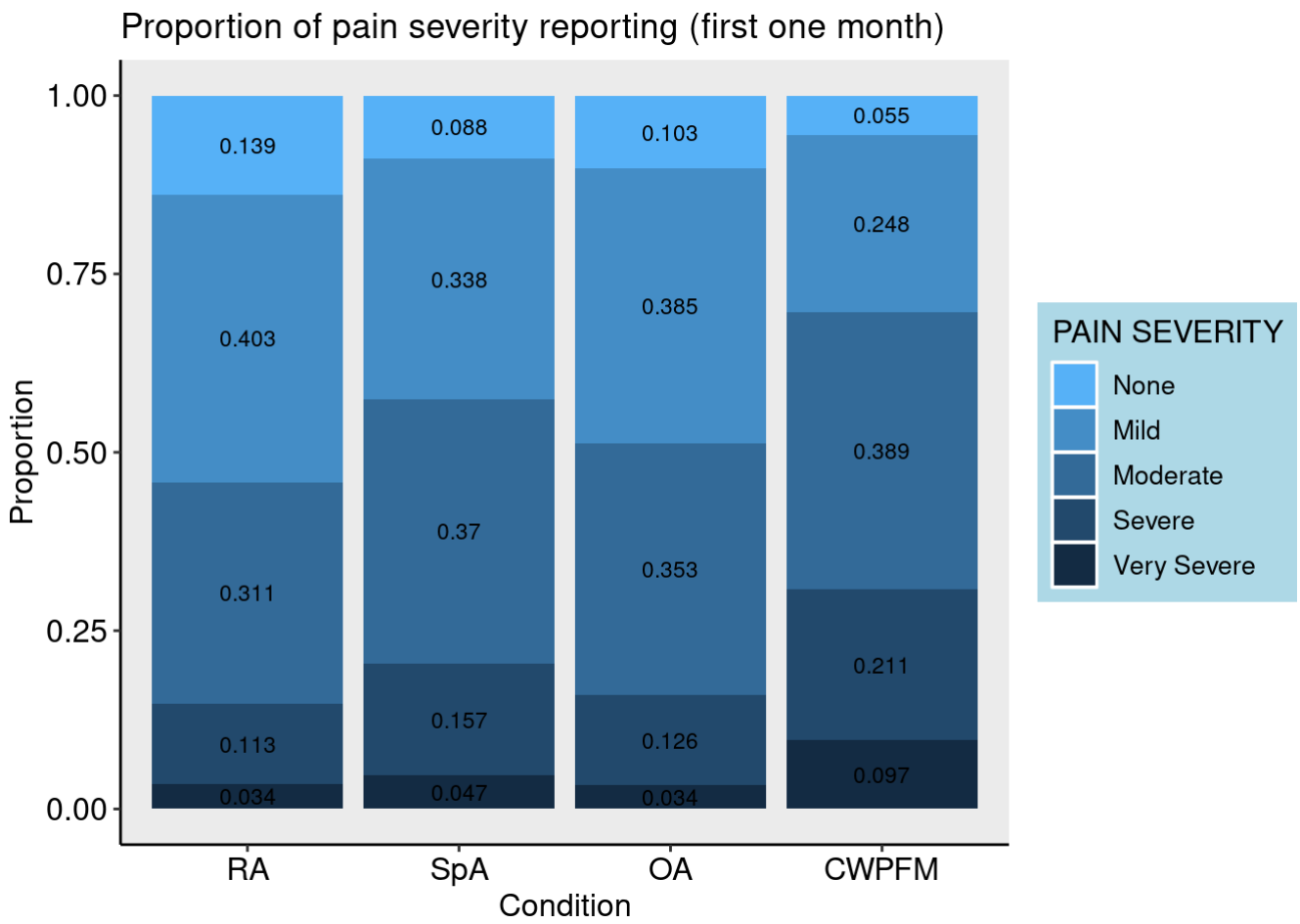
- BLUE - RA
- PURPLE - OA
- BLACK - CWPFM
- ORANGE - SpA

Code Output



PROPORTION OF PAIN LEVEL REPORTING FOR THE FIRST ONE MONTH

Code Output



MULTILEVEL MODEL

```
##Create new data name for the multilevel modelling and then factor the userid
#### RA ####
data.ra <- ra.painmotif
data.ra <- within(data.ra, {
  userid <- factor(userid)
})

#### SpA ####
data.spa <- spa.painmotif
data.spa <- within(data.spa, {
  userid <- factor(userid)
})

#### OA ####
data.oa <- oa.painmotif
data.oa <- within(data.oa, {
  userid <- factor(userid)
})

#### CWPFM ####
data.cwpfm <- cwpfm.painmotif
data.cwpfm <- within(data.cwpfm, {
  userid <- factor(userid)
})
```

Model 1: The Null Model

Equation: $painSeverity_{ij} = \beta_0 + u_{0j} + e_{ij}$

- $painSeverity_{ij}$ is the pain level at measurement occasion *i* (in the dataset, the measurement occasion is denoted by 'time') for individual *j* (in the dataset, the individual is denoted by 'userid').
- β_0 is the overall mean across individuals.
- u_{0j} is the random effect of individual *j* on pain level (Level 2). This is assumed to follow a normal distribution with mean zero and variance σ^2_{u0}.
- e_{ij} is the measurement occasion-level residual (Level 1).

Here in R, I use lme4 package for fitting in the mixed effect model (developed by Douglas Bates and Martin Maechler).

In the model,

- userid = blocking factor (random factor)
- time = within subject/repeated measures factor (fixed factor)
- painSeverity = measured (dependent) variable

Null model for RA

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ (1 | userid)
##    Data: data.ra
##
##      AIC      BIC   logLik deviance df.resid
##  19932.4  19953.5  -9963.2  19926.4     8387
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4799 -0.6501 -0.1236  0.6004  4.8134
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  userid   (Intercept) 0.4460   0.6678
##  Residual             0.5496   0.7414
## Number of obs: 8390, groups:  userid, 422
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  2.55041    0.03406   74.89
```

RESULT INTERPRETATION:

- The overall mean (across all individuals with RA only) is estimated as 2.55041.
- The mean for individual *j* is estimated as $2.55041 + \hat{u}_{0j}$, where $\hat{u}_{0j}$ is the individual residual which will be estimated in the next model. An individual with $\hat{u}_{0j} > 0$ has a mean that is higher than average pain, while $\hat{u}_{0j} < 0$ for an individual with below-average pain.
- The between-individual (level 2) variance 'userid (Intercept)' in pain severity is estimated as $\hat{\sigma}^2_{u0} = 0.4460$, and the within-individual (level 1) variance 'Residual' is estimated as $\hat{\sigma}^2_e = 0.5496$.

- The total variance is $0.4460 + 0.5496 = 0.9956$.
- The variance partition coefficient (VPC) is $0.4460/0.9956 = 0.45$, which indicates that 45% of the variance in pain severity can be attributed to differences between individuals with RA only.

Notes to self:

- both MLE and REML estimation procedure is based on optimizing a function of the log likelihood using penalized iteratively re-weighted least squares. The log likelihood is evaluated using an adaptive Gauss-Hermite approximation, which, when using the default value of one, reduces to the Laplacian approximation. This default approximation can be changed using the nAGQ=n option, where n is an integer >1, representing the number of points used for evaluating the adaptive Gauss-Hermite approximation. The greater the value of n, the more accurate the evaluation of the log likelihood, but it will take longer to fit the model.

Null model - testing for individual effects for RA

A likelihood ratio test (LRT) is carried out to compare Model 1 (as above) with a null single-level model (i.e. removing the random effect):

Equation: $painSeverity_{ij} = \beta_0 + e_{ij}$

Code Output

```
##
## Call:
## lm(formula = painSeverity ~ 1, data = data.ra)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.501 -0.501 -0.501  0.499  2.499
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.50095    0.01069   233.9   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9794 on 8389 degrees of freedom
```

```
## 'log Lik.' -9963.194 (df=3)
```

```
## 'log Lik.' -11730.05 (df=2)
```

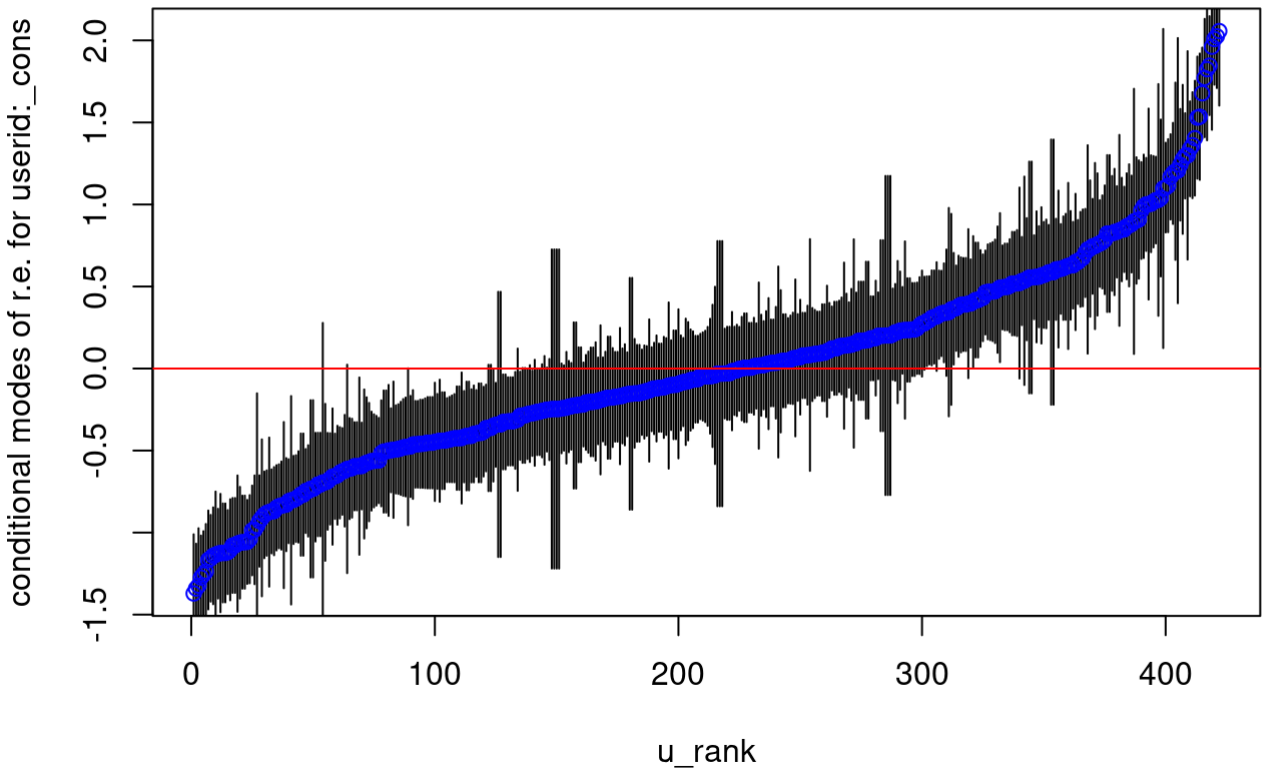
$LR = 2(-9963.194 - -11730.05) = 3533.712$ on 1 d.f.

By default, we know that 5% point of a chi-squared distribution on 1 d.f. is 3.84, there is strong evidence of indiividual effect on pain level and therefore, the multilevel model is more appropriate.

Null model - examining individual random effects (residuals) for RA

To estimate the individual-level residuals $\hat{u}_{0j}$ and their associated standard errors, [ranef] command is used with [postVar] option, creating a random effects object, which has variance-covariance matrix in the [postVar] attribute.

Code Output



Null model for OA

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ (1 | userid)
##    Data: data.oa
##
##      AIC      BIC   logLik deviance df.resid
##  19785.4  19806.5  -9889.7  19779.4     8380
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.2864 -0.6567 -0.1105  0.6362  4.6319
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  userid   (Intercept) 0.3684   0.6070
##  Residual             0.5465   0.7393
## Number of obs: 8383, groups:  userid, 408
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  2.61859    0.03156   82.96
```

RESULT INTERPRETATION:

- The overall mean (across all individuals with OA only) is estimated as 2.61859.
- The mean for individual *j* is estimated as $2.61859 + \hat{u}_{0j}$, where $\hat{u}_{0j}$ is the individual residual which will be estimated in the next model. An individual with $\hat{u}_{0j} > 0$ has a mean that is higher than average pain, while $\hat{u}_{0j} < 0$ for an individual with below-average pain.
- The between-individual (level 2) variance 'userid (Intercept)' in pain severity is estimated as $\hat{\sigma}^2_{u0} = 0.3684$, and the within-individual (level 1) variance 'Residual' is estimated as $\hat{\sigma}^2_e = 0.5465$.
- The total variance is $0.3684 + 0.5465 = 0.9149$.
- The variance partition coefficient (VPC) is $0.3684/0.9149 = 0.40$, which indicates that 40% of the variance in pain severity can be attributed to differences between individuals with OA only.

Null model testing for individual effects for OA

A likelihood ratio test (LRT) is carried out to compare Model 1 (as above) with a null single-level model (i.e. removing the random effect):

Equation: $painSeverity_{ij} = \beta_0 + e_{ij}$

Code Output

```
##
## Call:
## lm(formula = painSeverity ~ 1, data = data.oa)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6034 -0.6034  0.3966  0.3966  2.3966
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.60336    0.01035   251.4   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9481 on 8382 degrees of freedom
```

```
## 'log Lik.' -9889.698 (df=3)
```

```
## 'log Lik.' -11447.5 (df=2)
```

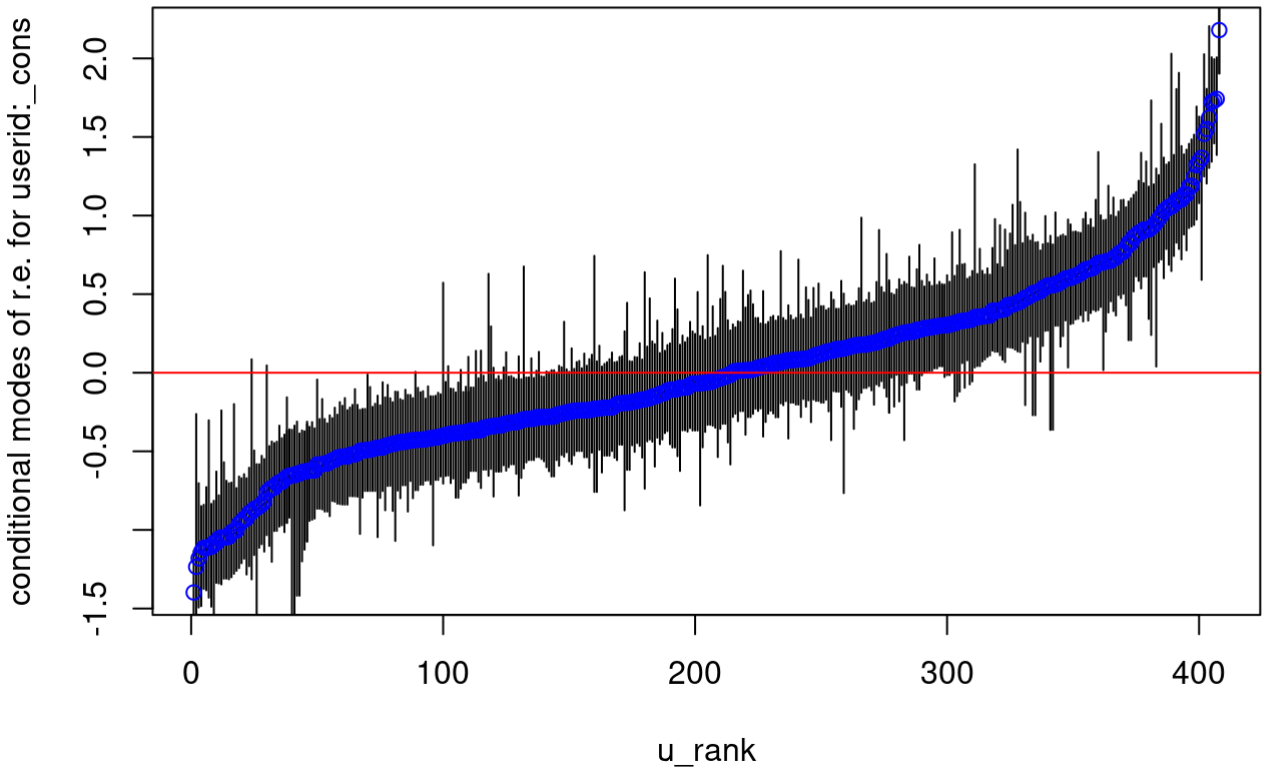
$LR = 2(-9889.698 - -11447.5) = 3115.604$ on 1 d.f.

By default, we know that 5% point of a chi-squared distribution on 1 d.f. is 3.84, there is strong evidence of indiivdual effect on pain level and therefore, the multilevel model is more appropriate.

Null model - examining individual random effects (residuals) for OA

Code Output

```
## Warning in u0tab.oa$u0 + 1.96 * u0tab$u0se: longer object length is not a
## multiple of shorter object length
```



Null model for CWPFM

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ (1 | userid)
##    Data: data.cwpfm
##
##      AIC      BIC   logLik deviance df.resid
##  11786.5  11806.0  -5890.3  11780.5     4789
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6364 -0.6778 -0.0110  0.6498  3.3606
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  userid   (Intercept) 0.4535   0.6734
##  Residual             0.5973   0.7729
## Number of obs: 4792, groups:  userid, 252
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  3.06631    0.04472   68.56
```

RESULT INTERPRETATION:

- The overall mean (across all individuals with RA only) is estimated as 3.06631.
- The mean for individual *j* is estimated as $3.06631 + \hat{u}_{0j}$, where $\hat{u}_{0j}$ is the individual residual which will be estimated in the next model. An individual with $\hat{u}_{0j} > 0$ has a mean that is higher than average pain, while $\hat{u}_{0j} < 0$ for an individual with below-average pain.
- The between-individual (level 2) variance 'userid (Intercept)' in pain severity is estimated as $\hat{\sigma}^2_{u0} = 0.4535$, and the within-individual (level 1) variance 'Residual' is estimated as $\hat{\sigma}^2_e = 0.5973$.
- The total variance is $0.4535 + 0.5973 = 1.0508$.
- The variance partition coefficient (VPC) is $0.4535/1,0508 = 0.43$, which indicates that 43% of the variance in pain severity can be attributed to differences between individuals with CWPFM only.

Null model - testing for individual effects for CWPFM

A likelihood ratio test (LRT) is carried out to compare Model 1 (as above) with a null single-level model (i.e. removing the random effect):

Equation: $painSeverity_{ij} = \beta_0 + e_{ij}$

Code Output

```
##
## Call:
## lm(formula = painSeverity ~ 1, data = data.cwpfm)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.04591 -1.04591 -0.04591  0.95409  1.95409
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.04591    0.01491   204.2   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.032 on 4791 degrees of freedom
```

```
## 'log Lik.' -5890.274 (df=3)
```

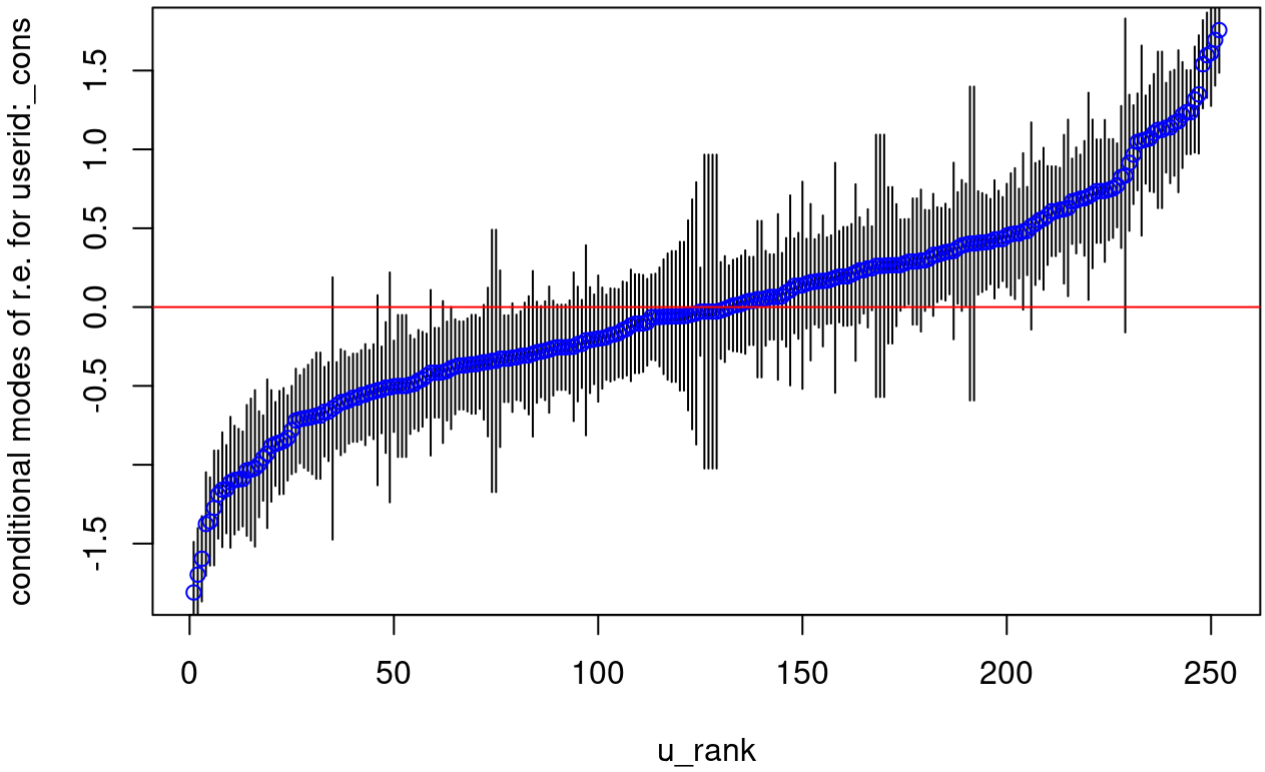
```
## 'log Lik.' -6951.577 (df=2)
```

$LR = 2(-5890.274 - -6951.577) = 2122.606$ on 1 d.f.

By default, we know that 5% point of a chi-squared distribution on 1 d.f. is 3.84, there is strong evidence of indiividual effect on pain level and therefore, the multilevel model is more appropriate.

Null model - examining individual random effects (residuals) for CWPFM

Code Output



Null model for SpA

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ (1 | userid)
##    Data: data.spa
##
##      AIC      BIC   logLik deviance df.resid
##   4802.2   4818.8  -2398.1   4796.2     1902
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.2814 -0.6739 -0.1003  0.6442  3.7615
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  userid   (Intercept) 0.3234   0.5687
##  Residual             0.6449   0.8030
## Number of obs: 1905, groups:  userid, 100
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  2.72369    0.06066    44.9
```

RESULT INTERPRETATION:

- The overall mean (across all individuals with SpA only) is estimated as 2.72369.
- The mean for individual *j* is estimated as $2.72369 + \hat{u}_{0j}$, where $\hat{u}_{0j}$ is the individual residual which will be estimated in the next model. An individual with $\hat{u}_{0j} > 0$ has a mean that is higher than average pain, while $\hat{u}_{0j} < 0$ for an individual with below-average pain.
- The between-individual (level 2) variance 'userid (Intercept)' in pain severity is estimated as $\hat{\sigma}^2_{u0} = 0.3234$, and the within-individual (level 1) variance 'Residual' is estimated as $\hat{\sigma}^2_e = 0.6449$.
- The total variance is $0.3234 + 0.6449 = 0.9683$.
- The variance partition coefficient (VPC) is $0.3234/0.9683 = 0.33$, which indicates that 33% of the variance in pain severity can be attributed to differences between individuals with SpA only.

Null model - testing for individual effects for SpA

A likelihood ratio test (LRT) is carried out to compare Model 1 (as above) with a null single-level model (i.e. removing the random effect):

Equation: $painSeverity_{ij} = \beta_0 + e_{ij}$

Code Output

```
##
## Call:
## lm(formula = painSeverity ~ 1, data = data.spa)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7365 -0.7365  0.2635  0.2635  2.2635
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.73648    0.02252   121.5   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.983 on 1904 degrees of freedom
```

```
## 'log Lik.' -2398.078 (df=3)
```

```
## 'log Lik.' -2669.863 (df=2)
```

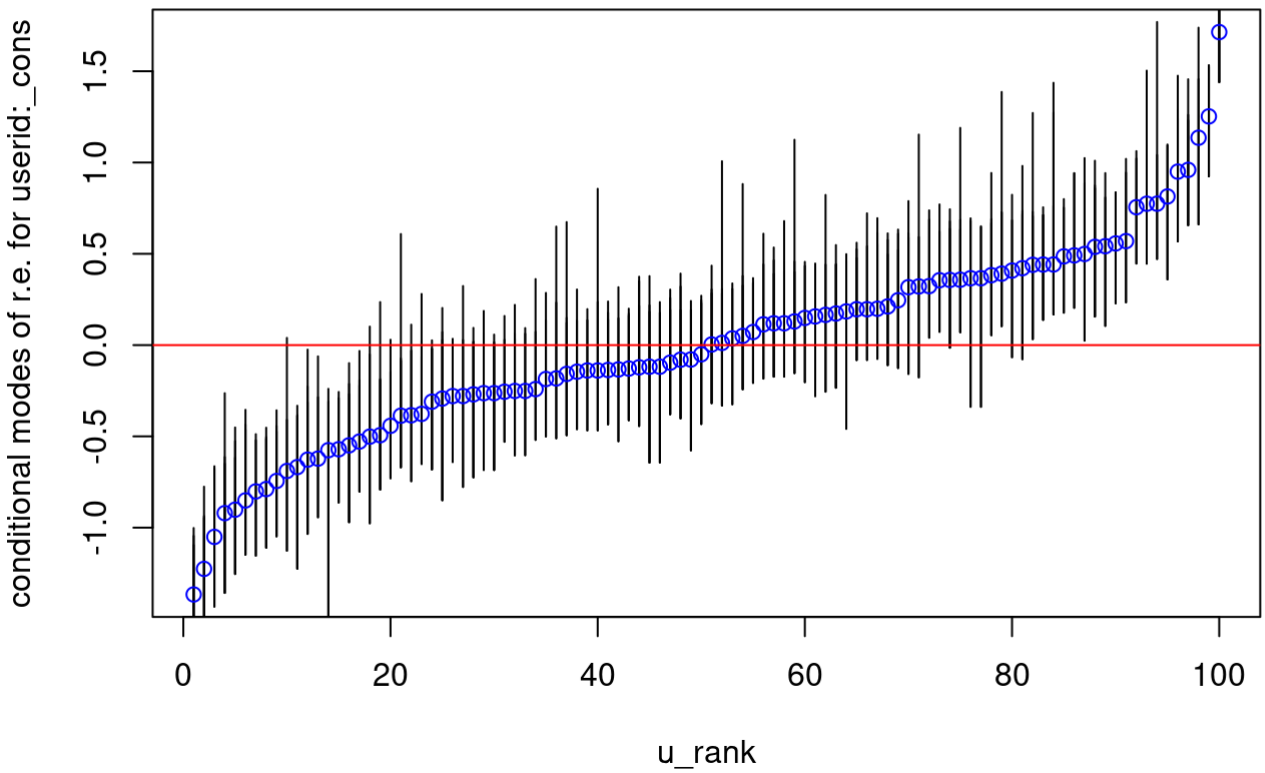
$LR = 2(-2398.078 - -2669.863) = 543.57$ on 1 d.f.

By default, we know that 5% point of a chi-squared distribution on 1 d.f. is 3.84, there is strong evidence of indiivdual effect on pain level and therefore, the multilevel model is more appropriate.

Null model- examining individual random effects (residuals) for SpA

Code Output

```
## Warning in u0tab.spa$u0 + 1.96 * u0tab$u0se: longer object length is not a
## multiple of shorter object length
```



Random intercept model for RA

Allowing for a linear effect,

Equation: $painSeverity_{ij} = \beta_0 + \beta_1 time_{ij} + u_{0j} + e_{ij}$

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ time + (1 | userid)
##    Data: data.ra
##
##      AIC      BIC   logLik deviance df.resid
##  19914.6  19942.7  -9953.3  19906.6     8386
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5328 -0.6454 -0.1181  0.6178  4.8022
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  userid   (Intercept) 0.4437   0.6661
##  Residual             0.5484   0.7405
## Number of obs: 8390, groups:  userid, 422
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  2.6055011  0.0361567  72.061
## time        -0.0041037  0.0009218  -4.452
##
## Correlation of Fixed Effects:
##      (Intr)
## time -0.342
```

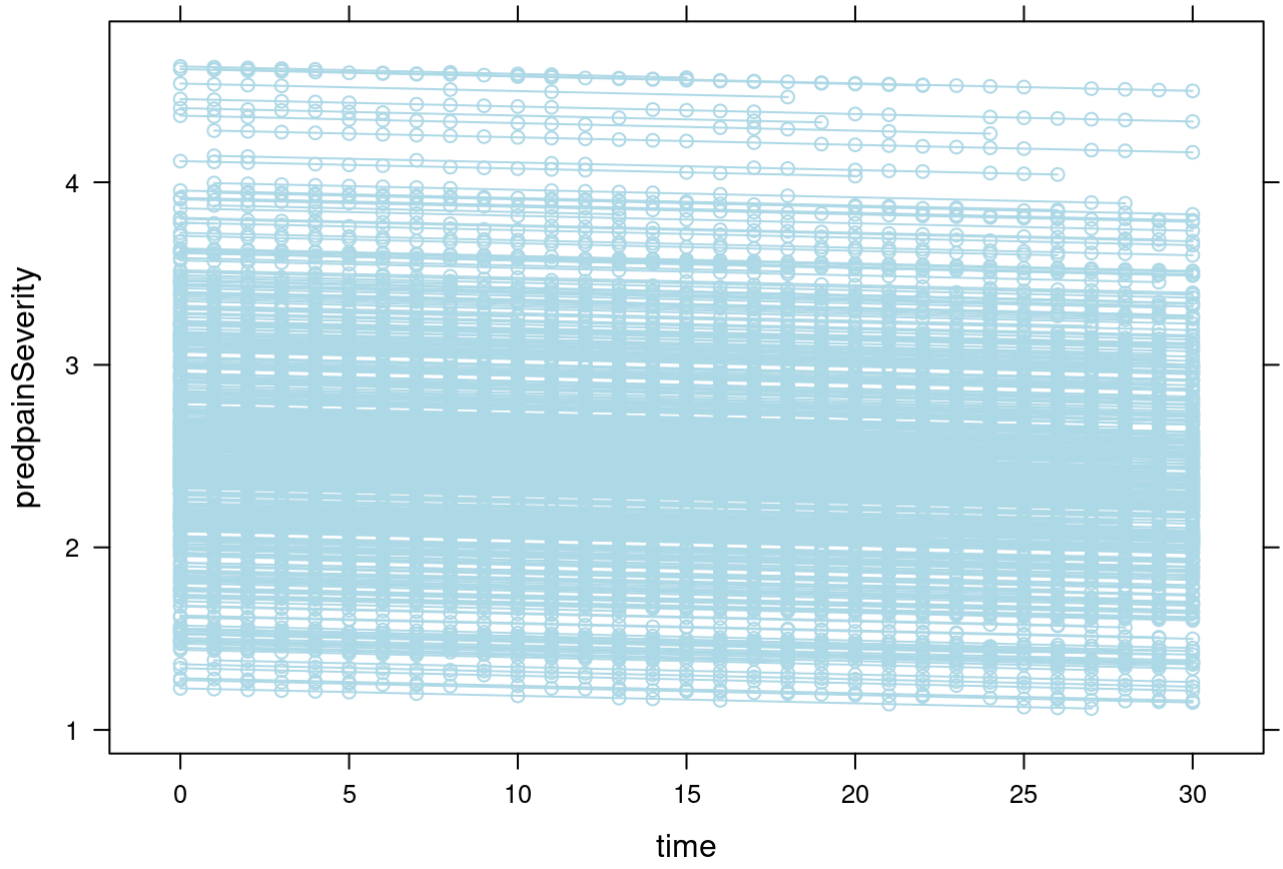
Result interpretation:

$$\hat{painSeverity}_{ij} = 2.606 - 0.0041037 time_{ij} + u_{0j} + e_{ij}$$

- The between-individual variance reduced slightly (compared to nullmodel) from 0.4460 to 0.4437 and the within-individual variance reduced slightly (compared to nullmodel) from 0.5496 to 0.5484.
- The total variance is $0.4437 + 0.5484 = 0.9921$ and the VPC is $0.4437/0.9921 = 0.4472$, which indicates that 45% of the variance in pain severity is attirbuted to differences between individuals.

Random intercept model - plot of the predicted individual lines for RA

Code Output



Random intercept model for OA

Allowing for a linear effect,

Equation: $painSeverity_{ij} = \beta_0 + \beta_1 time_{ij} + u_{0j} + e_{ij}$

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ time + (1 | userid)
##    Data: data.oa
##
##      AIC      BIC   logLik deviance df.resid
##  19777.4  19805.5  -9884.7  19769.4     8379
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.3398 -0.6443 -0.1112  0.6276  4.6250
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  userid   (Intercept) 0.3683   0.6069
##  Residual             0.5458   0.7388
## Number of obs: 8383, groups:  userid, 408
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  2.6581325  0.0339387  78.321
## time        -0.0028732  0.0009072  -3.167
##
## Correlation of Fixed Effects:
##      (Intr)
## time -0.368
```

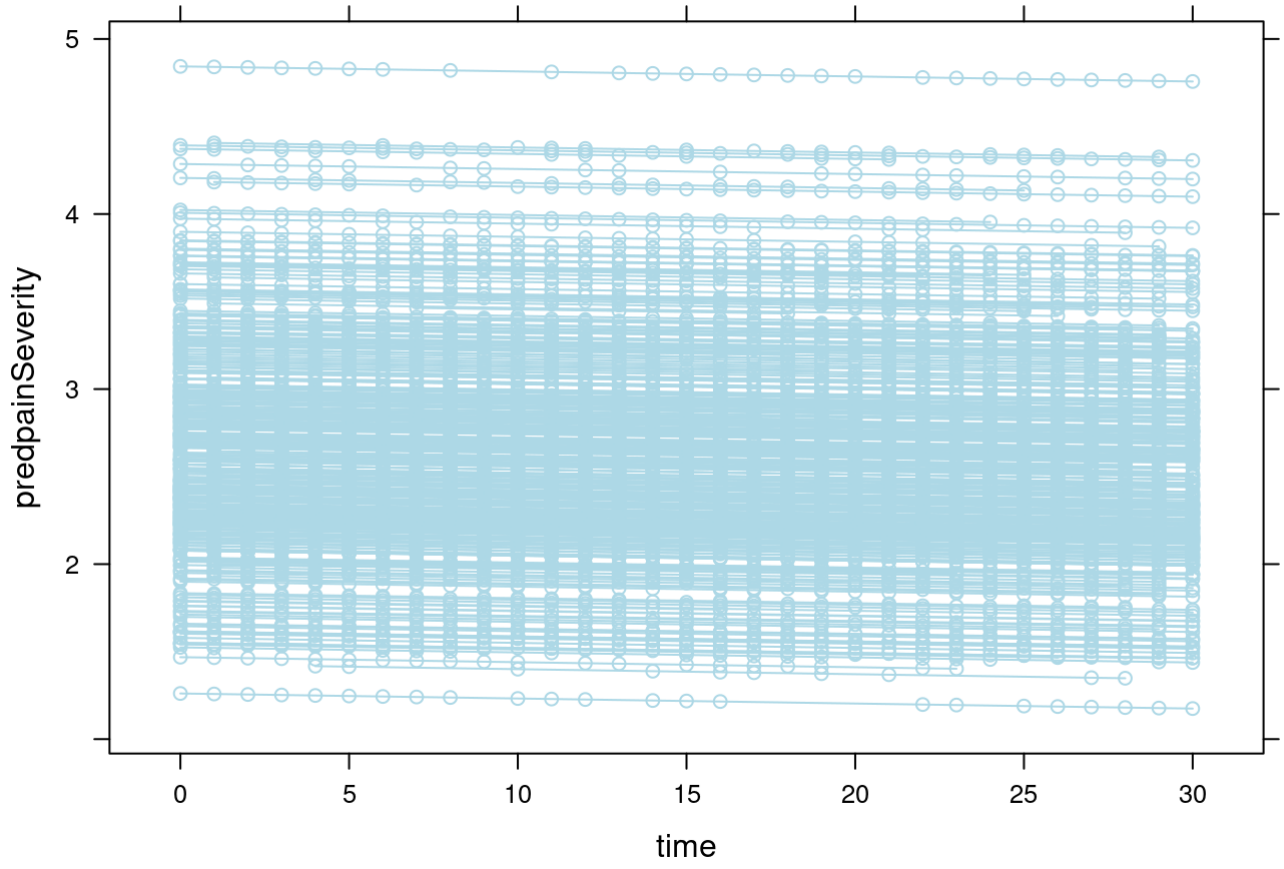
Result interpretation:

$$\hat{painSeverity}_{ij} = 2.658 - 0.0028732 time_{ij} + u_{0j} + e_{ij}$$

- The between-individual variance reduced minimally (compared to nullmodel) from 0.3684 to 0.3683 and the within-individual variance reduced slightly (compared to nullmodel) from 0.5465 to 0.5458.
- The total variance is $0.3683 + 0.5458 = 0.9141$ and the VPC is $0.3683/0.9141 = 0.4029$, which indicates that 41% of the variance in pain severity is attirbuted to differences between individuals.

Random intercept model - plot of the predicted individual lines for OA

Code Output



Random intercept model for CWPFM

Allowing for a linear effect,

Equation: $painSeverity_{ij} = \beta_0 + \beta_1 time_{ij} + u_{0j} + e_{ij}$

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ time + (1 | userid)
##    Data: data.cwpfm
##
##      AIC      BIC   logLik deviance df.resid
##  11778.8  11804.7  -5885.4  11770.8     4788
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6683 -0.6688 -0.0301  0.6341  3.3828
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  userid   (Intercept) 0.4526   0.6727
##  Residual             0.5961   0.7721
## Number of obs: 4792, groups:  userid, 252
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  3.117694   0.047620  65.470
## time        -0.003958   0.001270  -3.117
##
## Correlation of Fixed Effects:
##      (Intr)
## time -0.346
```

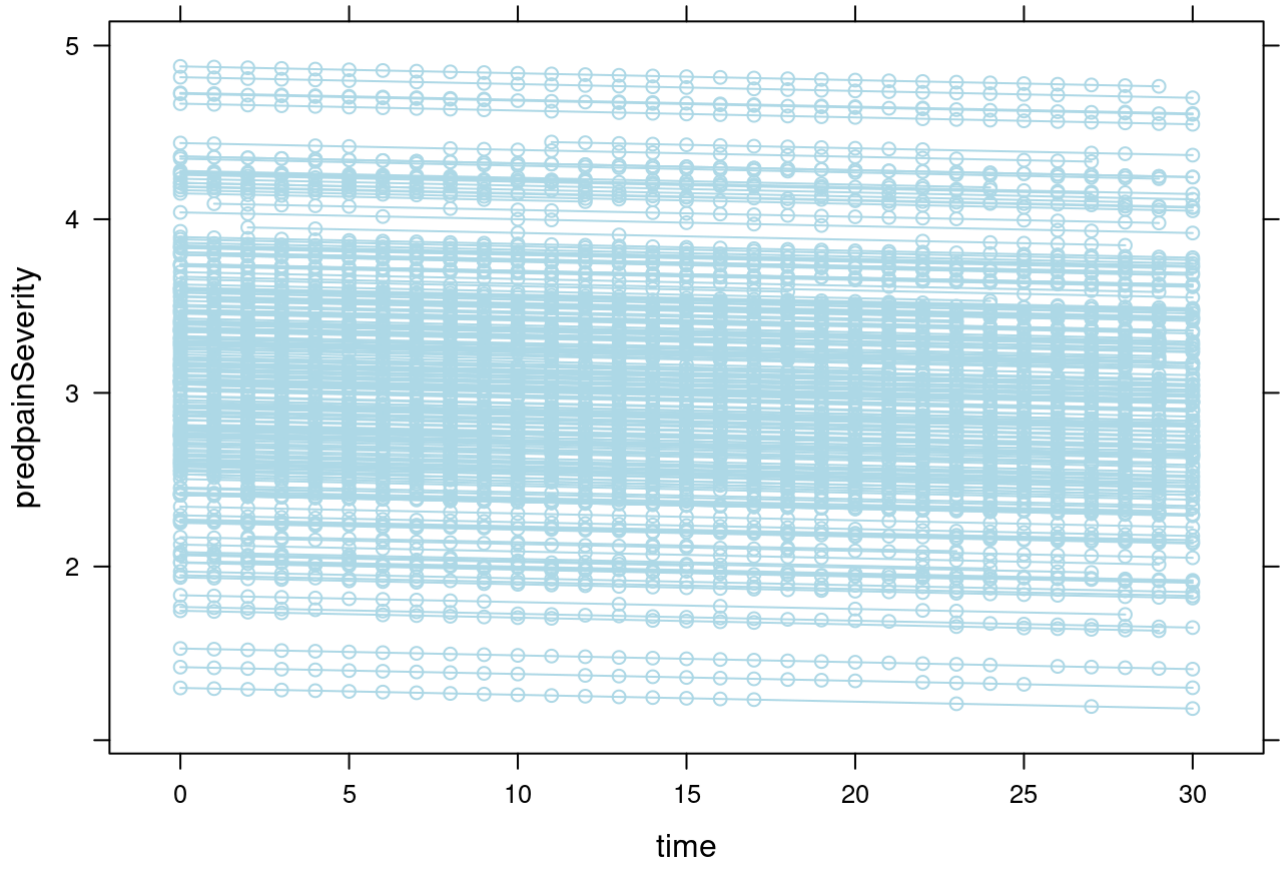
Result interpretation:

$$\hat{painSeverity}_{ij} = 3.118 - 0.003958 time_{ij} + u_{0j} + e_{ij}$$

- The between-individual variance reduced slightly (compared to nullmodel) from 0.4535 to 0.4526 and the within-individual variance reduced slightly (compared to nullmodel) from 0.5973 to 0.5961.
- The total variance is $0.4526 + 0.5961 = 1.0487$ and the VPC is $0.4526/1.0487 = 0.4472$, which indicates that 43% of the variance in pain severity is attirbuted to differences between individuals.

Random intercept model - plot of the predicted individual lines for CWPFM

Code Output



Random intercept model for SpA

Allowing for a linear effect,

Equation: $painSeverity_{ij} = \beta_0 + \beta_1 time_{ij} + u_{0j} + e_{ij}$

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ time + (1 | userid)
##    Data: data.spa
##
##      AIC      BIC   logLik deviance df.resid
##   4803.7   4825.9  -2397.9   4795.7     1901
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.3016 -0.6790 -0.0994  0.6404  3.7776
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  userid   (Intercept) 0.3227   0.5681
##  Residual             0.6448   0.8030
## Number of obs: 1905, groups:  userid, 100
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  2.741756   0.066606  41.164
## time        -0.001361   0.002083  -0.653
##
## Correlation of Fixed Effects:
##      (Intr)
## time -0.415
```

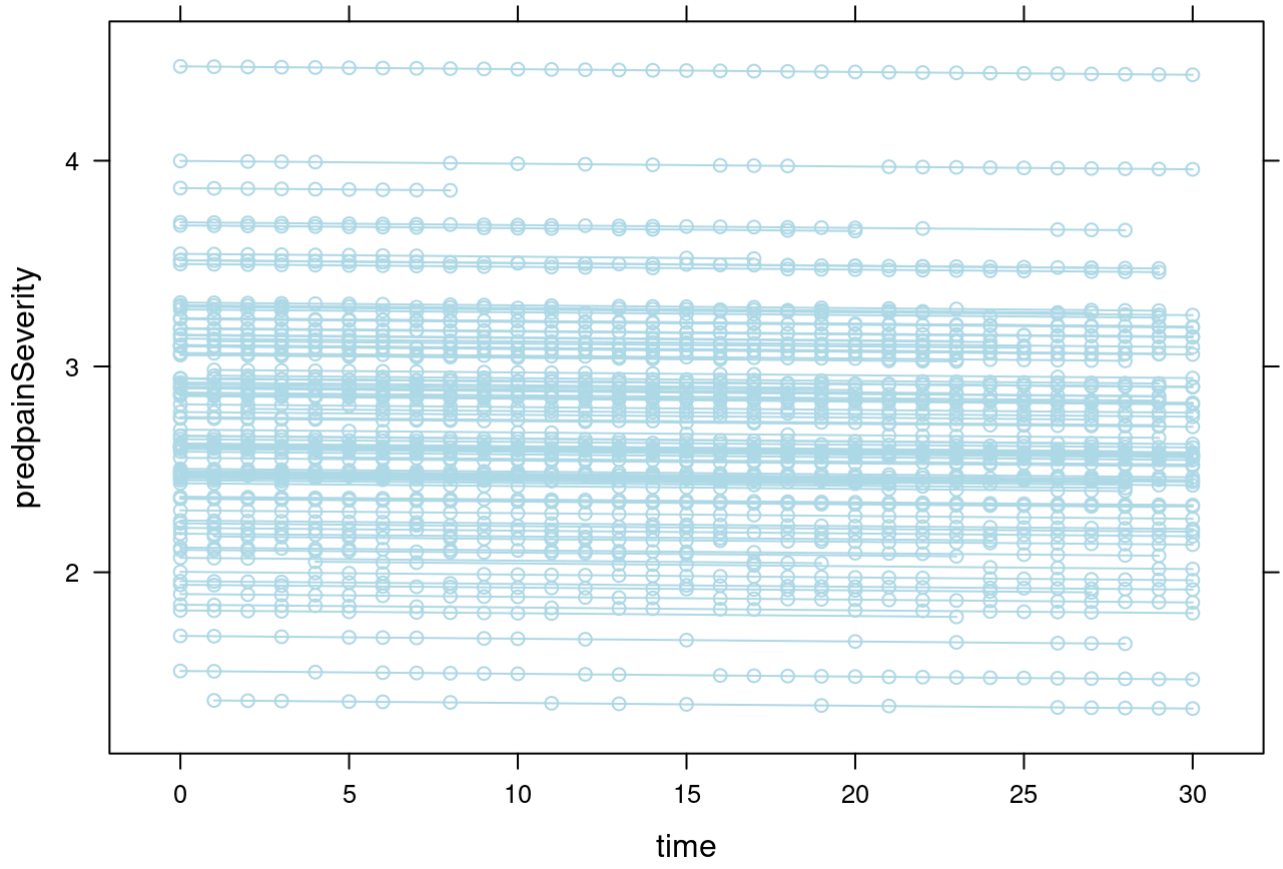
Result interpretation:

$$\hat{painSeverity}_{ij} = 2.742 - 0.001361 time_{ij} + u_{0j} + e_{ij}$$

- The between-individual variance reduced minimally (compared to nullmodel) from 0.3234 to 0.3227 and the within-individual variance reduced slightly (compared to nullmodel) from 0.6449 to 0.6448.
- The total variance is $0.3227 + 0.6448 = 0.9675$ and the VPC is $0.3227/0.9675 = 0.334$, which indicates that 33% of the variance in pain severity is attirbuted to differences between individuals.

Random intercept model - plot of the predicted individual lines for SpA

Code Output



Random slope model for RA

Allowing for a linear effect,

Equation: $painSeverity_{ij} = \beta_0 + \beta_1 time_{ij} + u_{0j} + u_{1j} time_{ij} + e_{ij}$

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ time + (1 + time | userid)
##    Data: data.ra
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##  19702.7  19744.9  -9845.3  19690.7     8384
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5094 -0.6472 -0.1188  0.6113  4.9614
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  userid   (Intercept) 0.5150158 0.71765
##           time        0.0004575 0.02139  -0.37
##  Residual             0.5113829 0.71511
## Number of obs: 8390, groups:  userid, 422
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  2.602686   0.038396  67.785
## time        -0.003929   0.001433  -2.742
##
## Correlation of Fixed Effects:
##      (Intr)
## time -0.463
```

```
## Data: data.ra
## Models:
## interceptmodel.ra: painSeverity ~ time + (1 | userid)
## slopemodel1.ra: painSeverity ~ time + (1 + time | userid)
##                   Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## interceptmodel.ra  4 19915 19943 -9953.3    19907
## slopemodel1.ra     6 19703 19745 -9845.3    19691 215.92      2  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##              (Intercept)          time
## (Intercept)  0.515015846 -0.0056868932
## time        -0.005686893  0.0004575286
## attr(,"stddev")
## (Intercept)        time
##  0.71764605  0.02138992
## attr(,"correlation")
##             (Intercept)       time
## (Intercept)   1.0000000 -0.3704722
## time         -0.3704722  1.0000000
```

```
## Computing profile confidence intervals ...
```

```
##                    2.5 %       97.5 %
## .sig01       0.661820994  0.779368067
## .sig02      -0.485513674 -0.240423653
## .sig03       0.018892125  0.024081553
## .sigma       0.703885024  0.726636673
## (Intercept)  2.527310126  2.678196183
## time        -0.006747991 -0.001107273
```

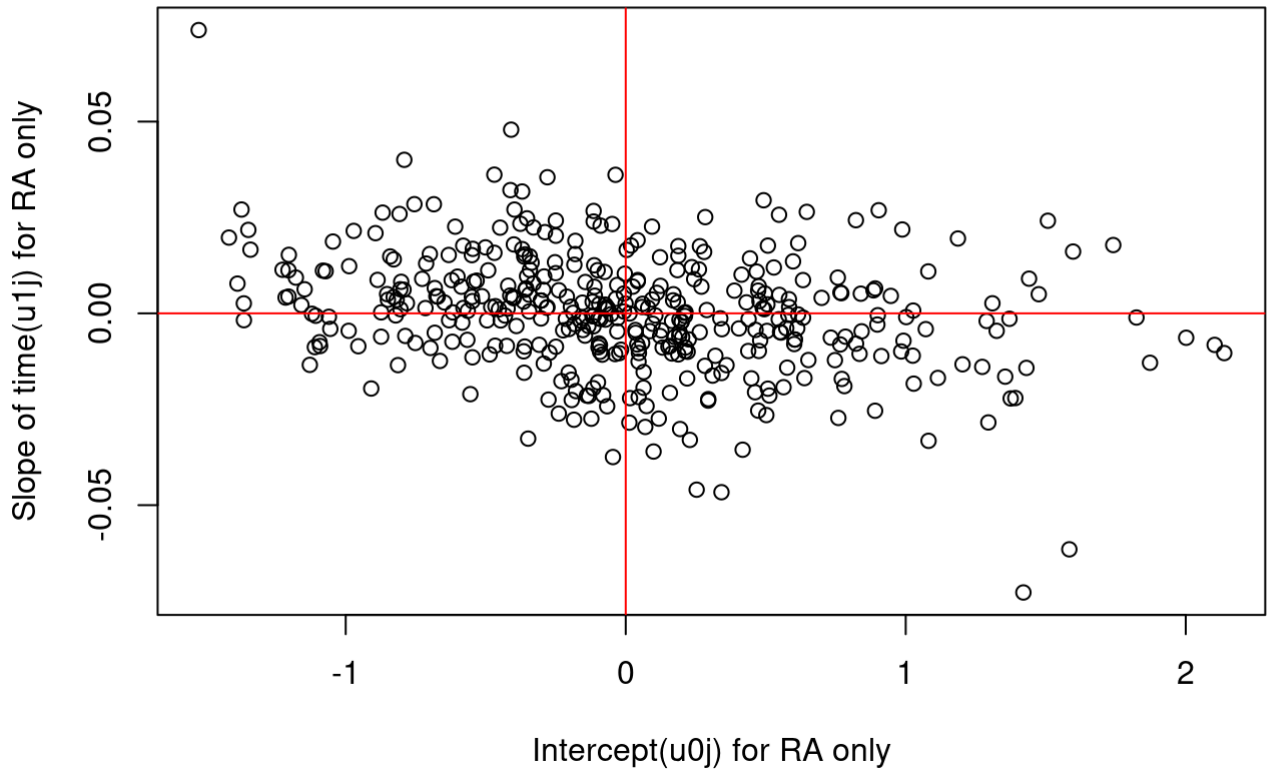
RESULT INTERPRETATION:

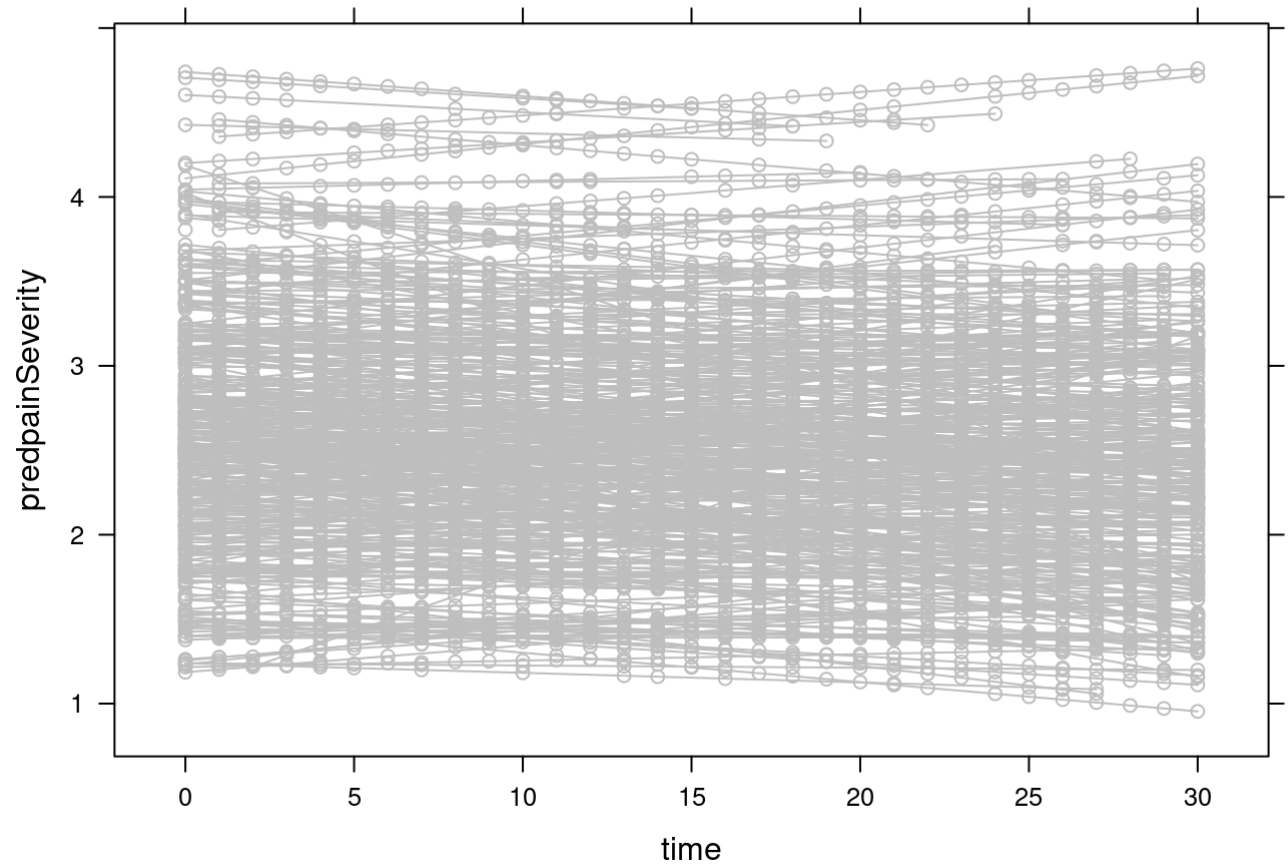
$$painSeverity_{ij} = 2.603 - 0.003929time_{ij} + u_{0j} + 0.0004575time_{ij} + e_{ij}$$

The ANOVA informs us that the random slope model is a better fit.

Random slope model plots for RA

Code Output





Random slope model for OA

Allowing for a linear effect,

Equation: $painSeverity_{ij} = \beta_0 + \beta_1 time_{ij} + u_{0j} + u_{1j} time_{ij} + e_{ij}$

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ time + (1 + time | userid)
##    Data: data.oa
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##  19636.0  19678.2  -9812.0  19624.0     8377
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.1049 -0.6323 -0.0882  0.6192  4.7277
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  userid   (Intercept) 0.4282985 0.65445
##           time        0.0003586 0.01894  -0.37
##  Residual             0.5165306 0.71870
## Number of obs: 8383, groups:  userid, 408
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  2.657298   0.036037   73.74
## time        -0.002717   0.001332   -2.04
##
## Correlation of Fixed Effects:
##      (Intr)
## time -0.479
```

```
## Data: data.oa
## Models:
## interceptmodel.oa: painSeverity ~ time + (1 | userid)
## slopemodel1.oa: painSeverity ~ time + (1 + time | userid)
##                   Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## interceptmodel.oa  4 19777 19806 -9884.7    19769
## slopemodel1.oa     6 19636 19678 -9812.0    19624 145.37      2  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##             (Intercept)          time
## (Intercept)  0.428298482 -0.0045741762
## time        -0.004574176  0.0003585689
## attr(,"stddev")
## (Intercept)        time
##  0.65444517  0.01893592
## attr(,"correlation")
##             (Intercept)       time
## (Intercept)   1.0000000 -0.3691079
## time         -0.3691079  1.0000000
```

```
## Computing profile confidence intervals ...
```

```
##                    2.5 %        97.5 %
## .sig01       0.601100859  7.134791e-01
## .sig02      -0.487721957 -2.338178e-01
## .sig03       0.016443507  2.158791e-02
## .sigma       0.707413107  7.302901e-01
## (Intercept)  2.586539772  2.728161e+00
## time        -0.005334858 -9.685741e-05
```

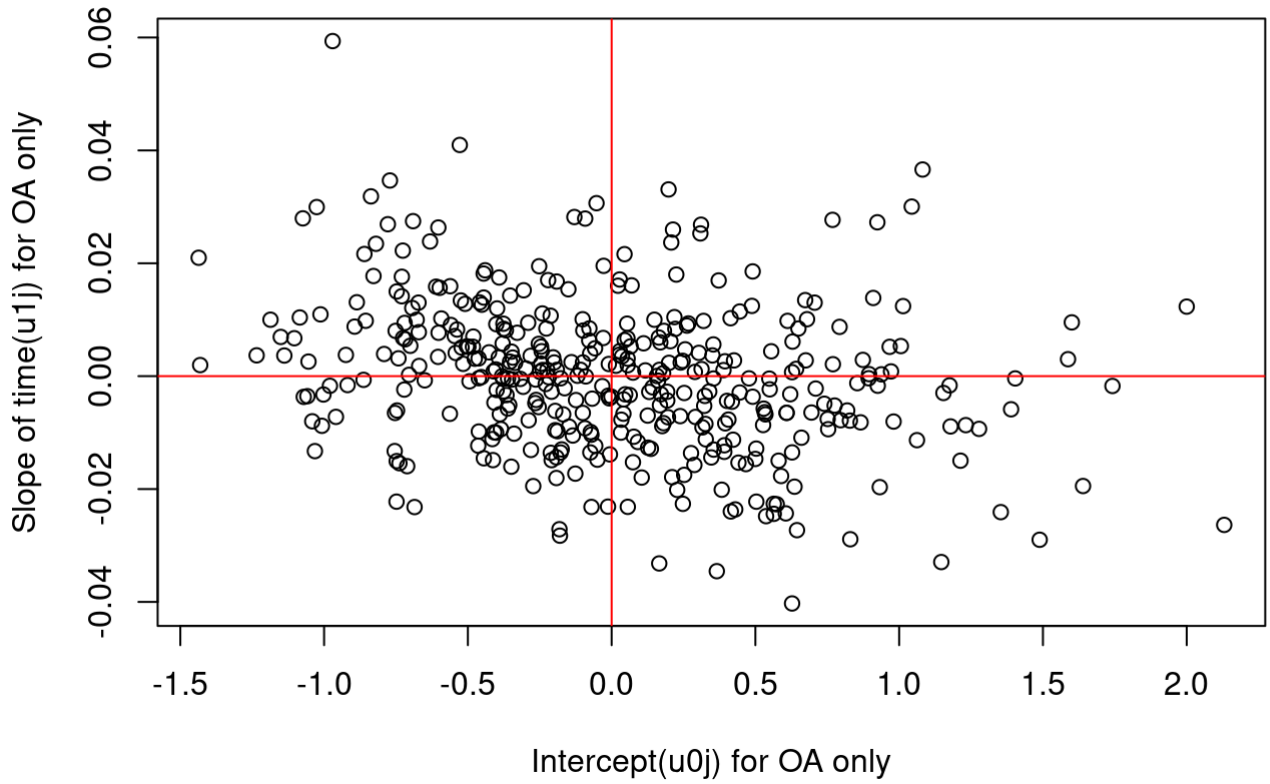
RESULT INTERPRETATION:

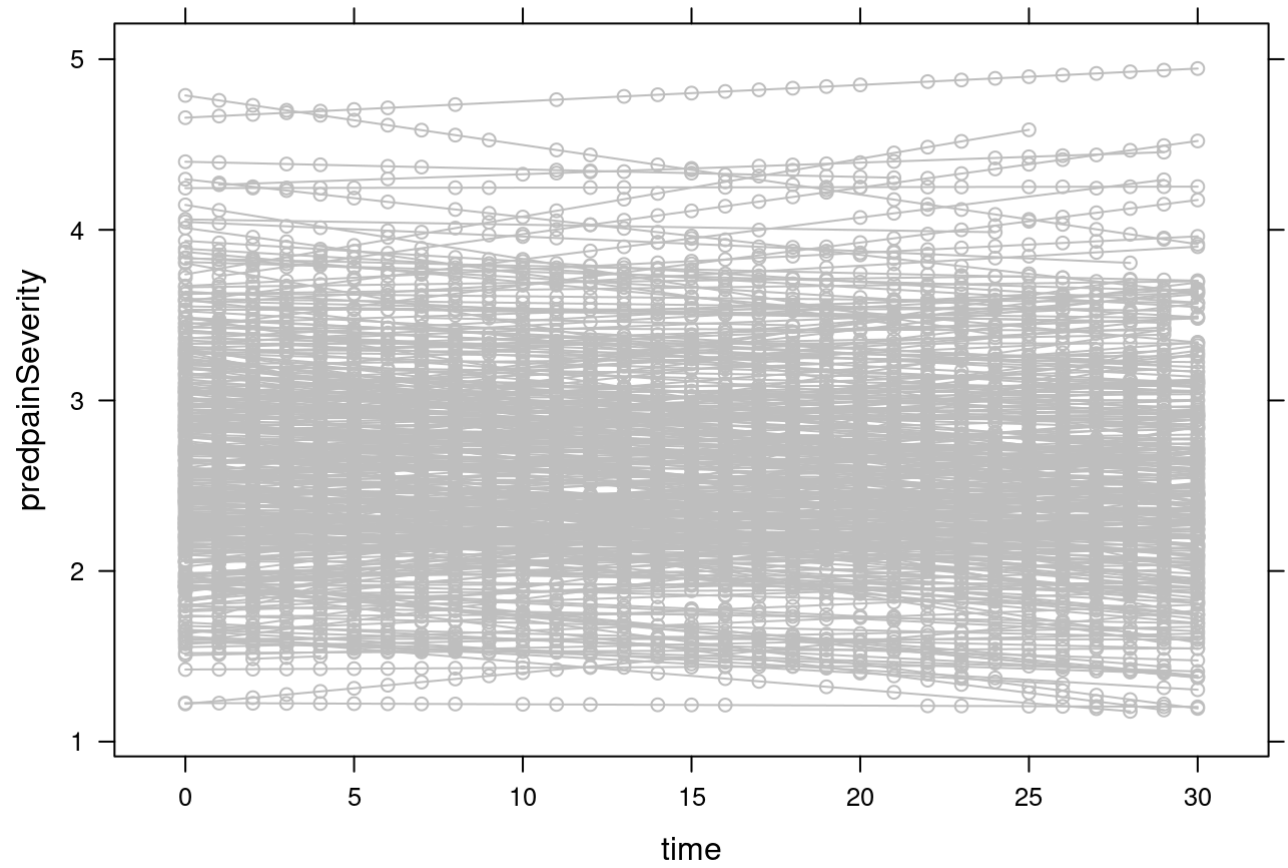
$$painSeverity_{ij} = 2.657 - 0.002717time_{ij} + u_{0j} + 0.0003586time_{ij} + e_{ij}$$

The ANOVA informs us that the random slope model is a better fit.

Random slope model plots for OA

Code Output





Random slope model for CWPFM

Allowing for a linear effect,

Equation: $painSeverity_{ij} = \beta_0 + \beta_1 time_{ij} + u_{0j} + u_{1j} time_{ij} + e_{ij}$

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ time + (1 + time | userid)
##    Data: data.cwpfm
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##  11701.5  11740.4  -5844.8  11689.5     4786
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6492 -0.6621 -0.0354  0.6323  3.6311
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  userid   (Intercept) 0.4489956 0.67007
##           time        0.0003884 0.01971  -0.18
##  Residual             0.5654470 0.75196
## Number of obs: 4792, groups:  userid, 252
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  3.117147   0.047390  65.776
## time        -0.004089   0.001845  -2.216
##
## Correlation of Fixed Effects:
##      (Intr)
## time -0.357
```

```
## Data: data.cwpfm
## Models:
## interceptmodel.cwpfm: painSeverity ~ time + (1 | userid)
## slopemodel1.cwpfm: painSeverity ~ time + (1 + time | userid)
##                      Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## interceptmodel.cwpfm  4 11779 11805 -5885.4    11771
## slopemodel1.cwpfm     6 11702 11740 -5844.8    11690 81.293      2  < 2.2e-16
##
## interceptmodel.cwpfm
## slopemodel1.cwpfm    ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##              (Intercept)          time
## (Intercept)  0.448995628 -0.0023744857
## time        -0.002374486  0.0003884453
## attr(,"stddev")
## (Intercept)        time
##  0.67007136  0.01970901
## attr(,"correlation")
##             (Intercept)       time
## (Intercept)   1.0000000 -0.1797975
## time         -0.1797975  1.0000000
```

```
## Computing profile confidence intervals ...
```

```
##                    2.5 %        97.5 %
## .sig01       0.601917160  0.7475045087
## .sig02      -0.358194536  0.0218432409
## .sig03       0.016182583  0.0235316375
## .sigma       0.736380775  0.7680928174
## (Intercept)  3.023961098  3.2104363703
## time        -0.007726408 -0.0004583978
```

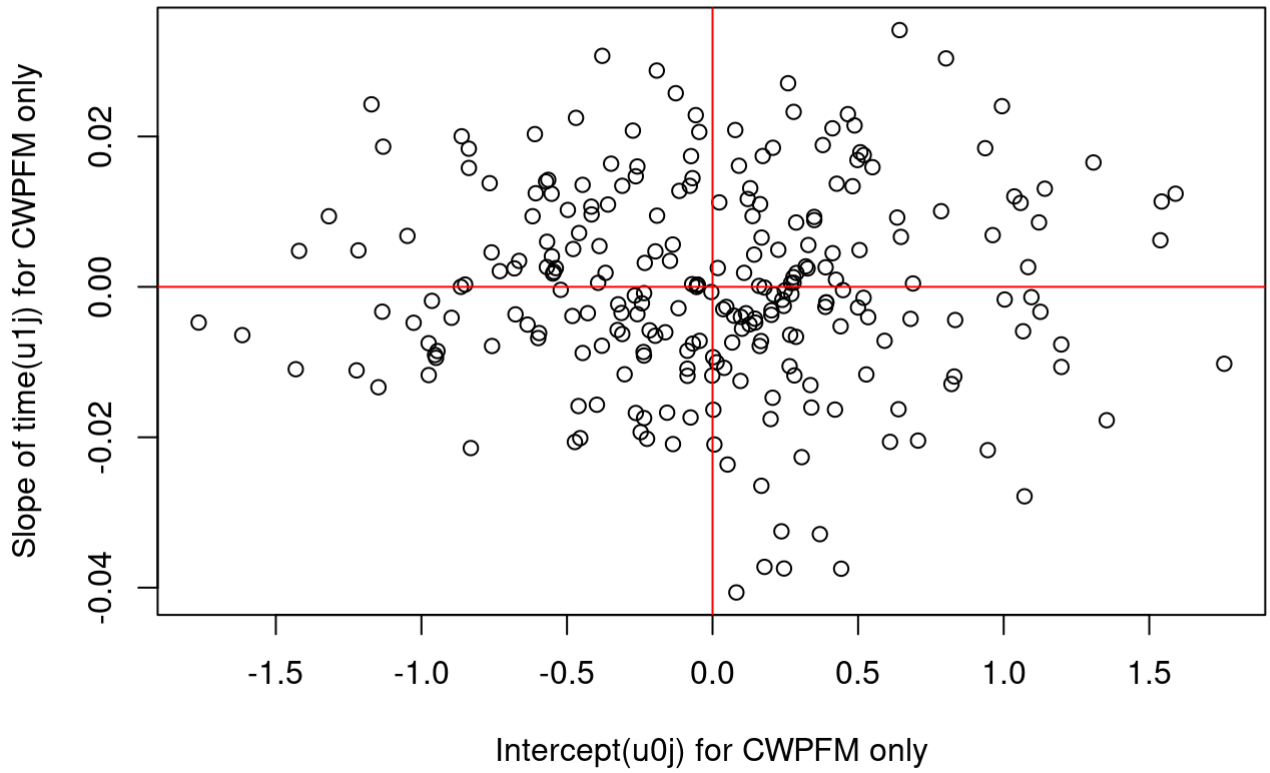
RESULT INTERPRETATION:

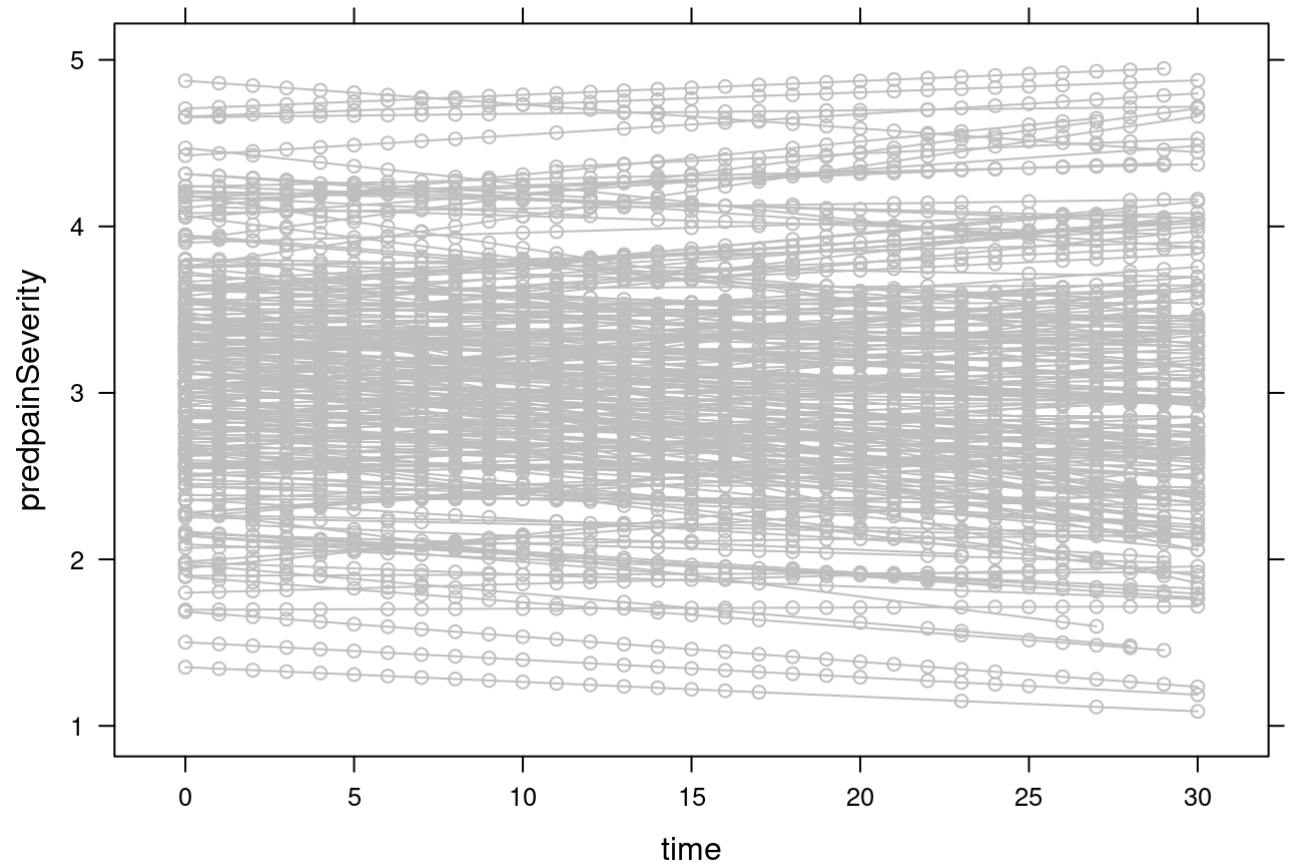
$$painSeverity_{ij} = 3.117 - 0.004089time_{ij} + u_{0j} + 0.0003884time_{ij} + e_{ij}$$

The ANOVA informs us that the random slope model is a better fit.

Random slope model plots for CWPFM

Code Output





Random slope model for SpA

Allowing for a linear effect,

Equation: $painSeverity_{ij} = \beta_0 + \beta_1 time_{ij} + u_{0j} + u_{1j} time_{ij} + e_{ij}$

Code Output

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ time + (1 + time | userid)
##    Data: data.spa
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   4745.2   4778.5  -2366.6   4733.2     1899
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.1674 -0.6677 -0.0606  0.5988  3.9387
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  userid   (Intercept) 0.4454702 0.66744
##           time        0.0006156 0.02481  -0.54
##  Residual             0.5939061 0.77065
## Number of obs: 1905, groups:  userid, 100
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  2.736439   0.075044   36.46
## time        -0.000726   0.003307   -0.22
##
## Correlation of Fixed Effects:
##      (Intr)
## time -0.598
```

```
## Data: data.spa
## Models:
## interceptmodel.spa: painSeverity ~ time + (1 | userid)
## slopemodel1.spa: painSeverity ~ time + (1 + time | userid)
##                    Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## interceptmodel.spa  4 4803.7 4825.9 -2397.9   4795.7
## slopemodel1.spa     6 4745.2 4778.5 -2366.6   4733.2 62.508      2   2.67e-14
##
## interceptmodel.spa
## slopemodel1.spa    ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##              (Intercept)          time
## (Intercept)  0.445470168 -0.0089094953
## time        -0.008909495  0.0006156385
## attr(,"stddev")
## (Intercept)        time
##  0.66743552  0.02481206
## attr(,"correlation")
##             (Intercept)       time
## (Intercept)   1.0000000 -0.5379983
## time         -0.5379983  1.0000000
```

```
## Computing profile confidence intervals ...
```

```
##                    2.5 %       97.5 %
## .sig01       0.56165230  0.796553768
## .sig02      -0.70584974 -0.304572983
## .sig03       0.01936901  0.031202487
## .sigma       0.74343747  0.797158642
## (Intercept)  2.58764126  2.884699007
## time        -0.00725753  0.005848157
```

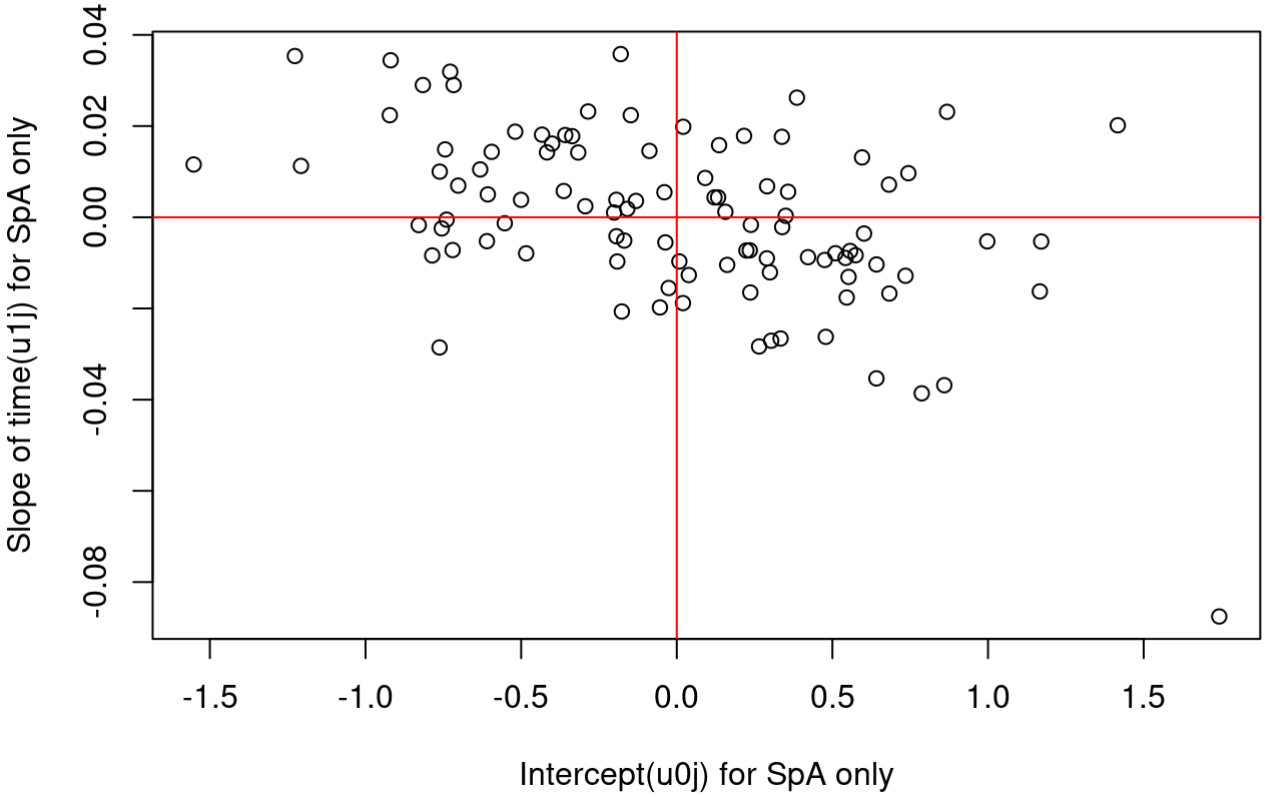
RESULT INTERPRETATION:

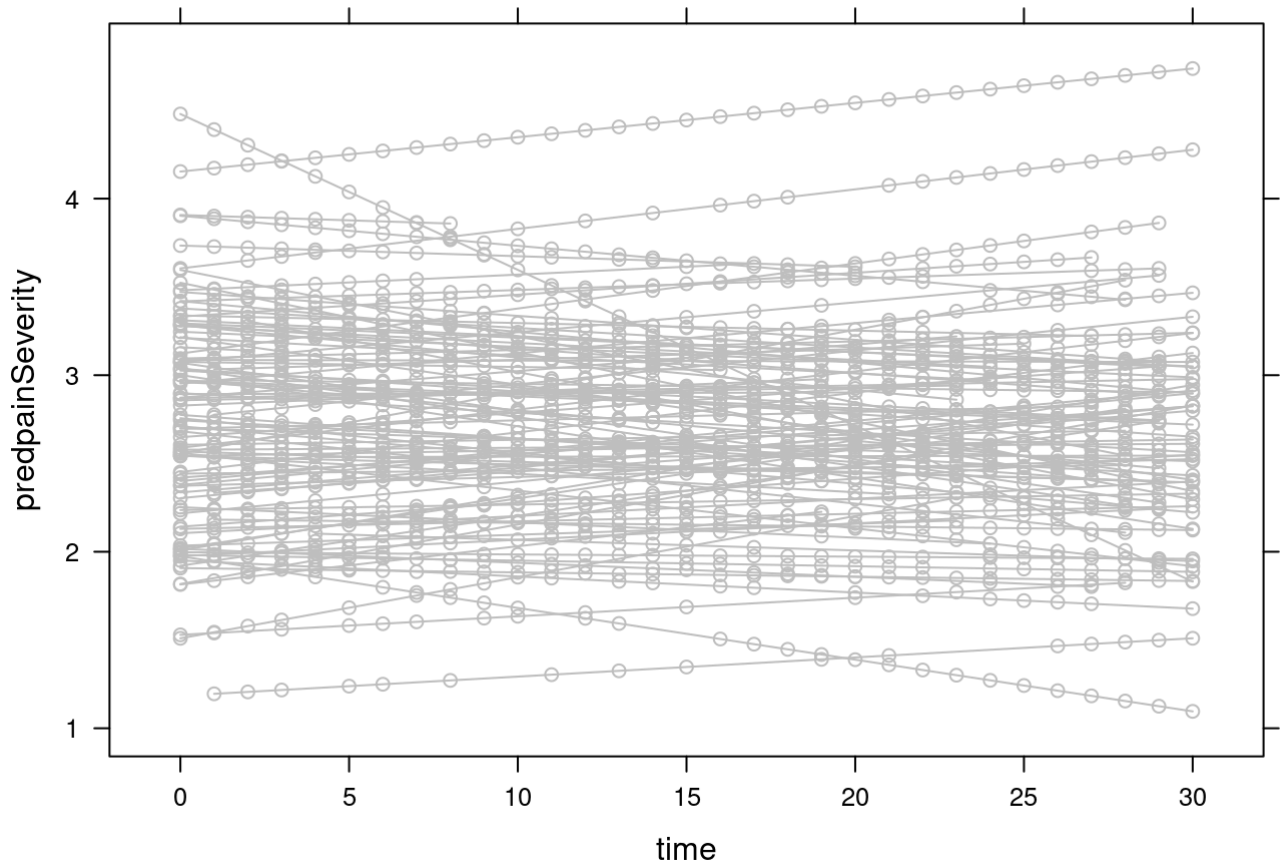
$$painSeverity_{ij} = 2.736 - 0.000726time_{ij} + u_{0j} + 0.0006156time_{ij} + e_{ij}$$

The ANOVA informs us that the random slope model is a better fit.

Random slope model plots for SpA

Code Output


Slope of time(u1j) for SpA only
0.04
0.02
0.00
-0.04
-0.08
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
Intercept(u0j) for SpA only

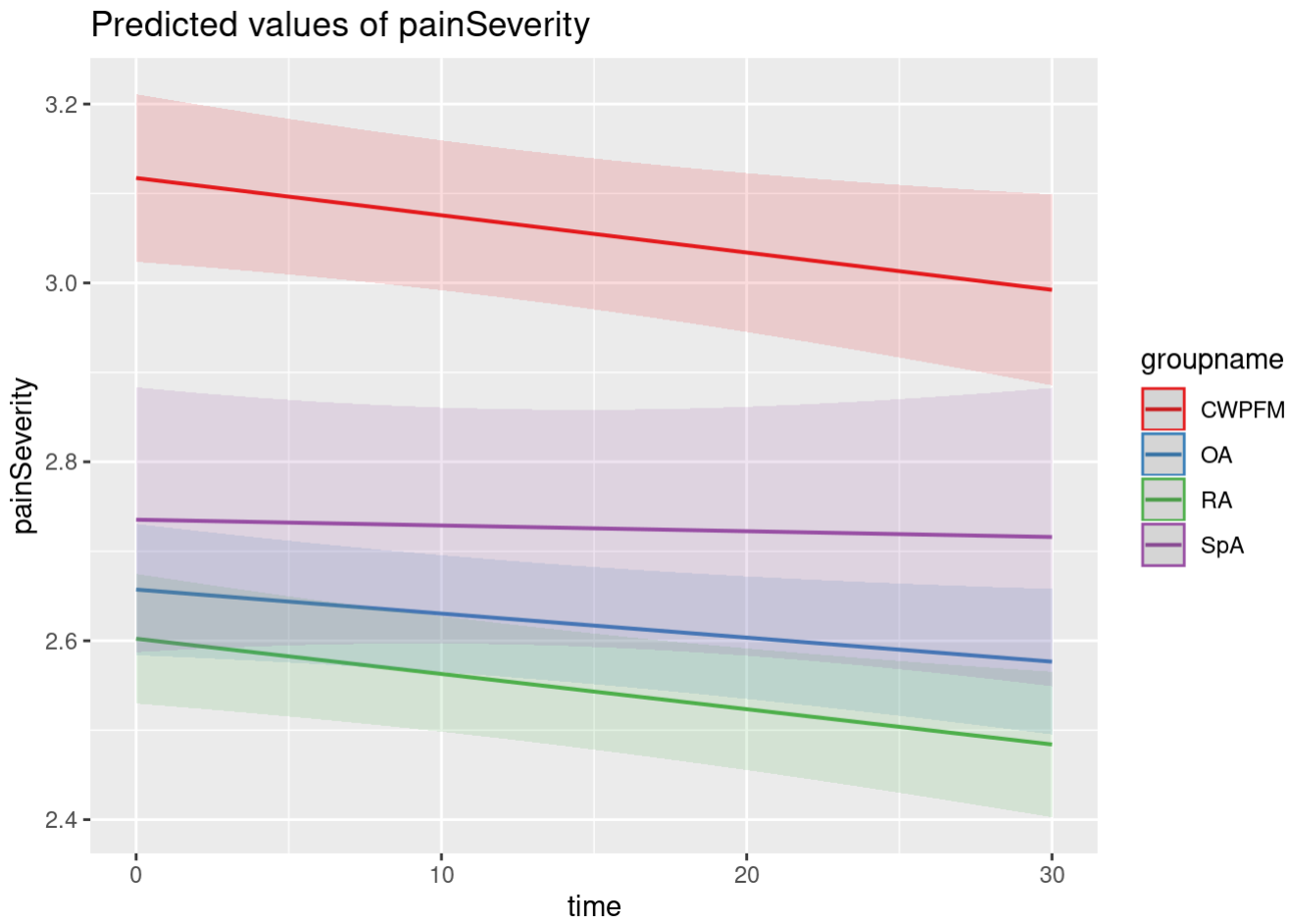

predpainSeverity
4
3
2
1
0
5
10
15
20
25
30
time

Predicted value plot for all diseases (random slope model)

Code Output

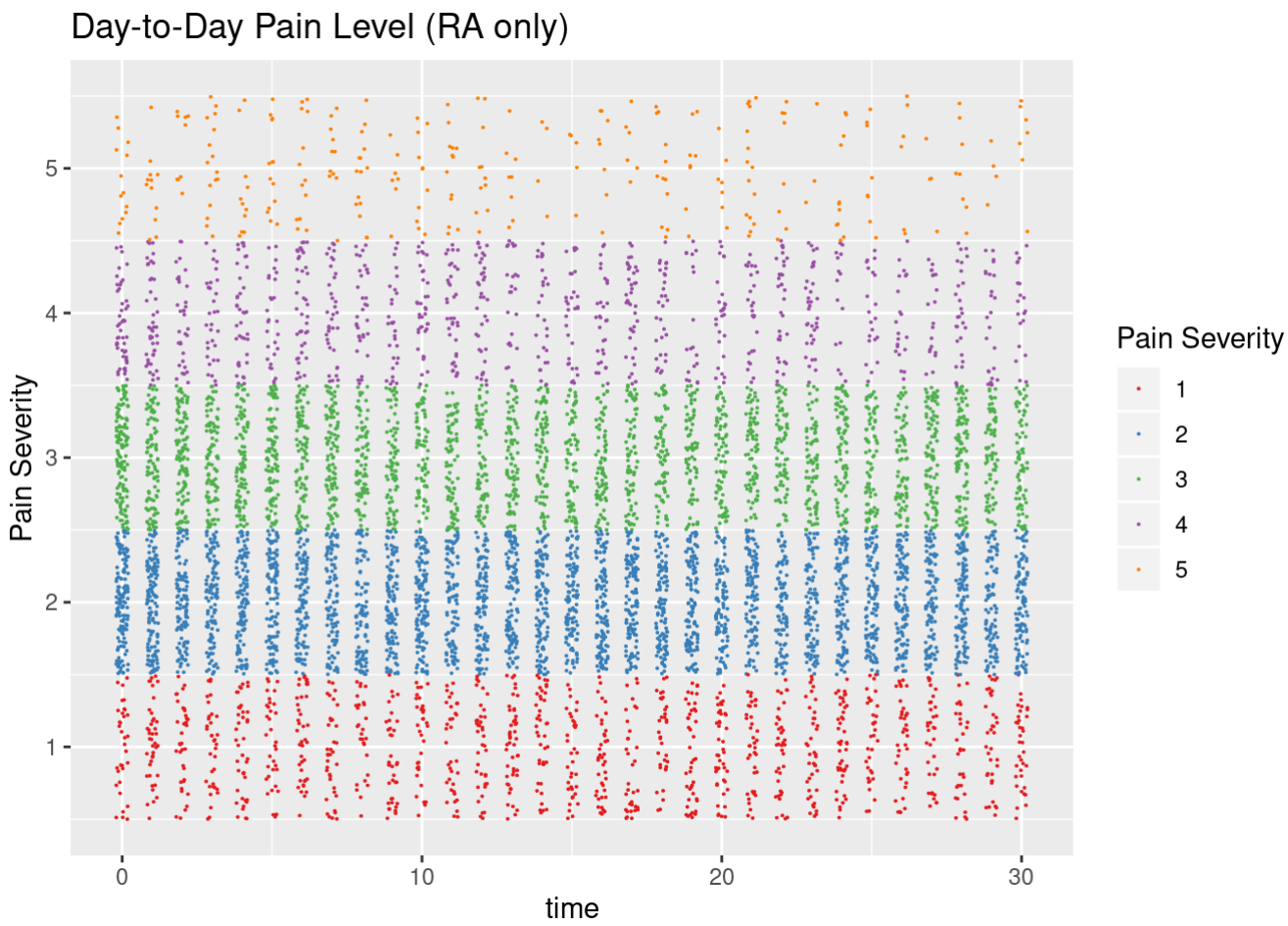
```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Mod
el is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: painSeverity ~ time + groupname + time * groupname + (1 + time |
##     userid)
##    Data: combo.all
## Control: lmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##  55804.7  55901.5 -27890.3  55780.7    23458
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4529 -0.6454 -0.0889  0.6198  4.8616
##
## Random effects:
##  Groups   Name        Variance  Std.Dev. Corr
##  userid   (Intercept) 0.4649080 0.68184
##           time        0.0004219 0.02054  -0.35
##  Residual             0.5308825 0.72862
## Number of obs: 23470, groups:  userid, 1182
##
## Fixed effects:
##                    Estimate Std. Error t value
## (Intercept)        3.117304   0.047802  65.212
## time              -0.004168   0.001854  -2.248
## groupnameOA       -0.460104   0.060677  -7.583
## groupnameRA       -0.515065   0.060397  -8.528
## groupnameSpA      -0.382016   0.089341  -4.276
## time:groupnameOA   0.001482   0.002325   0.637
## time:groupnameRA   0.000228   0.002330   0.098
## time:groupnameSpA  0.003524   0.003446   1.022
##
## Correlation of Fixed Effects:
##             (Intr) time   grpnOA grpnRA grpnSA tm:gOA tm:gRA
## time        -0.454
## groupnameOA -0.788  0.358
## groupnameRA -0.791  0.359  0.624
## groupnamSpA -0.535  0.243  0.422  0.423
## tim:grpnmOA  0.362 -0.797 -0.456 -0.286 -0.194
## tim:grpnmRA  0.361 -0.796 -0.285 -0.456 -0.193  0.635
## tm:grpnmSpA  0.244 -0.538 -0.192 -0.193 -0.458  0.429  0.428
## convergence code: 0
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```



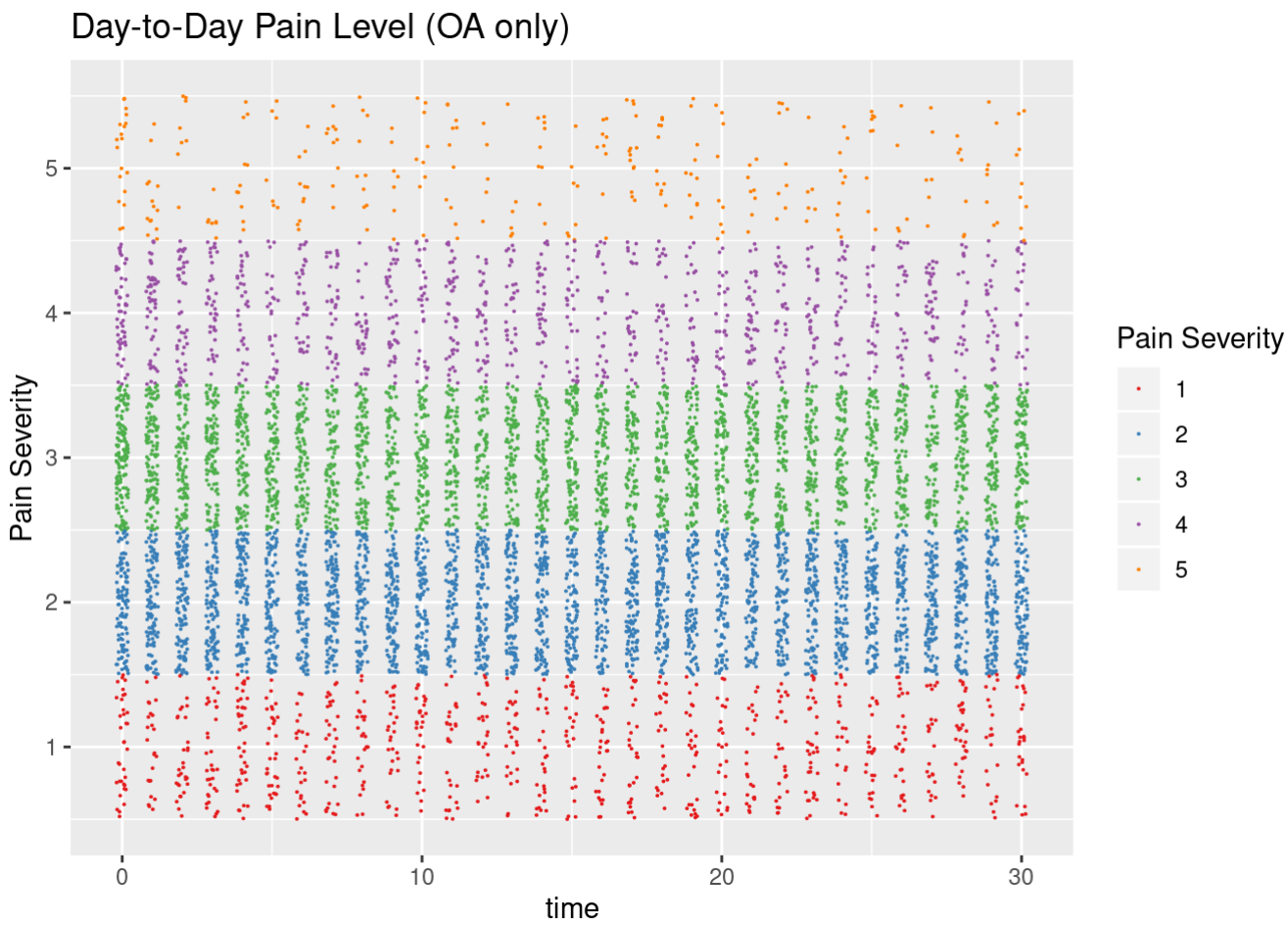
Day-to-day pain scoring for RA (jitter plot)

Code Output



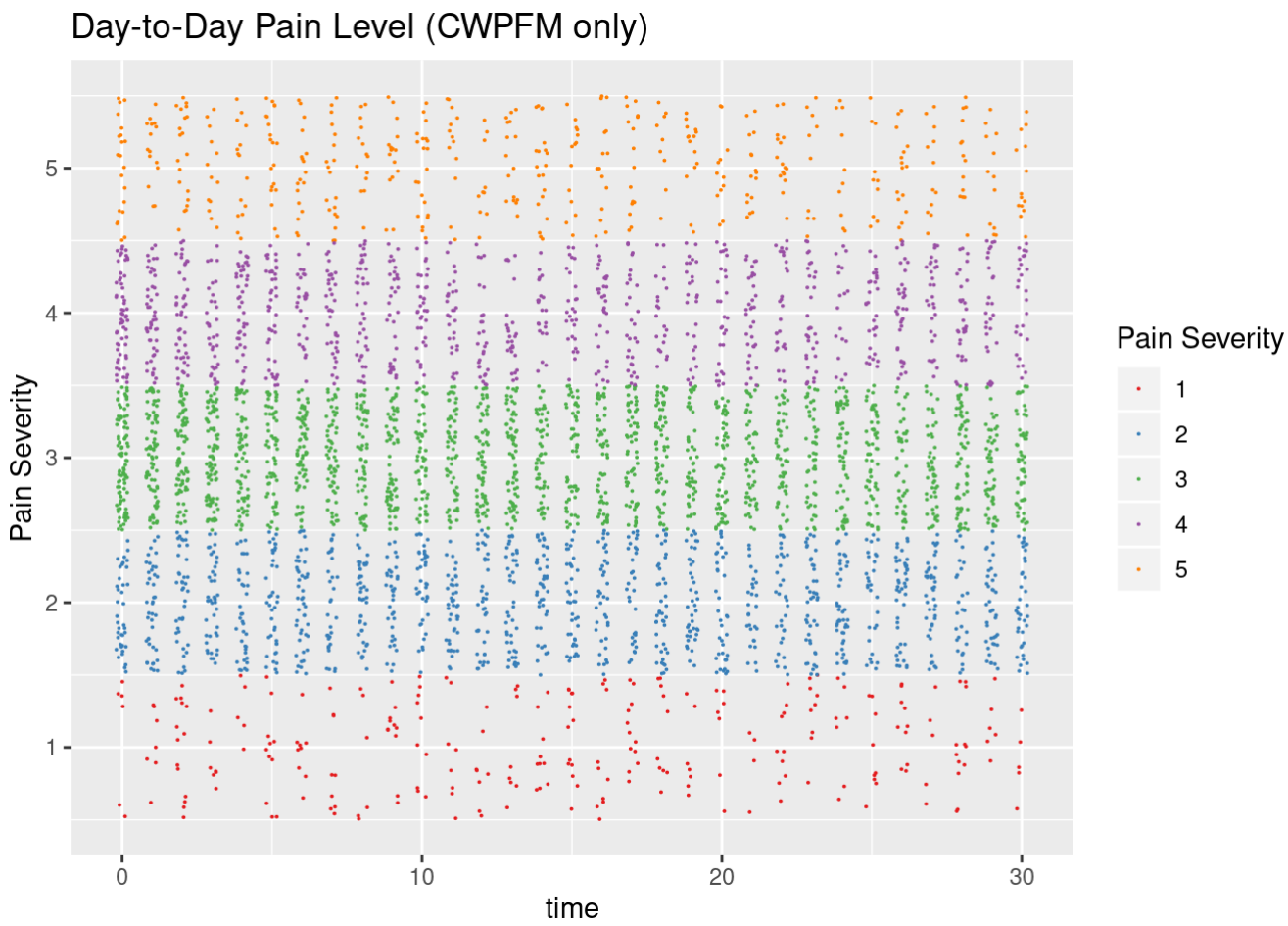
Day-to-day pain scoring for OA (jitter plot)

Code Output



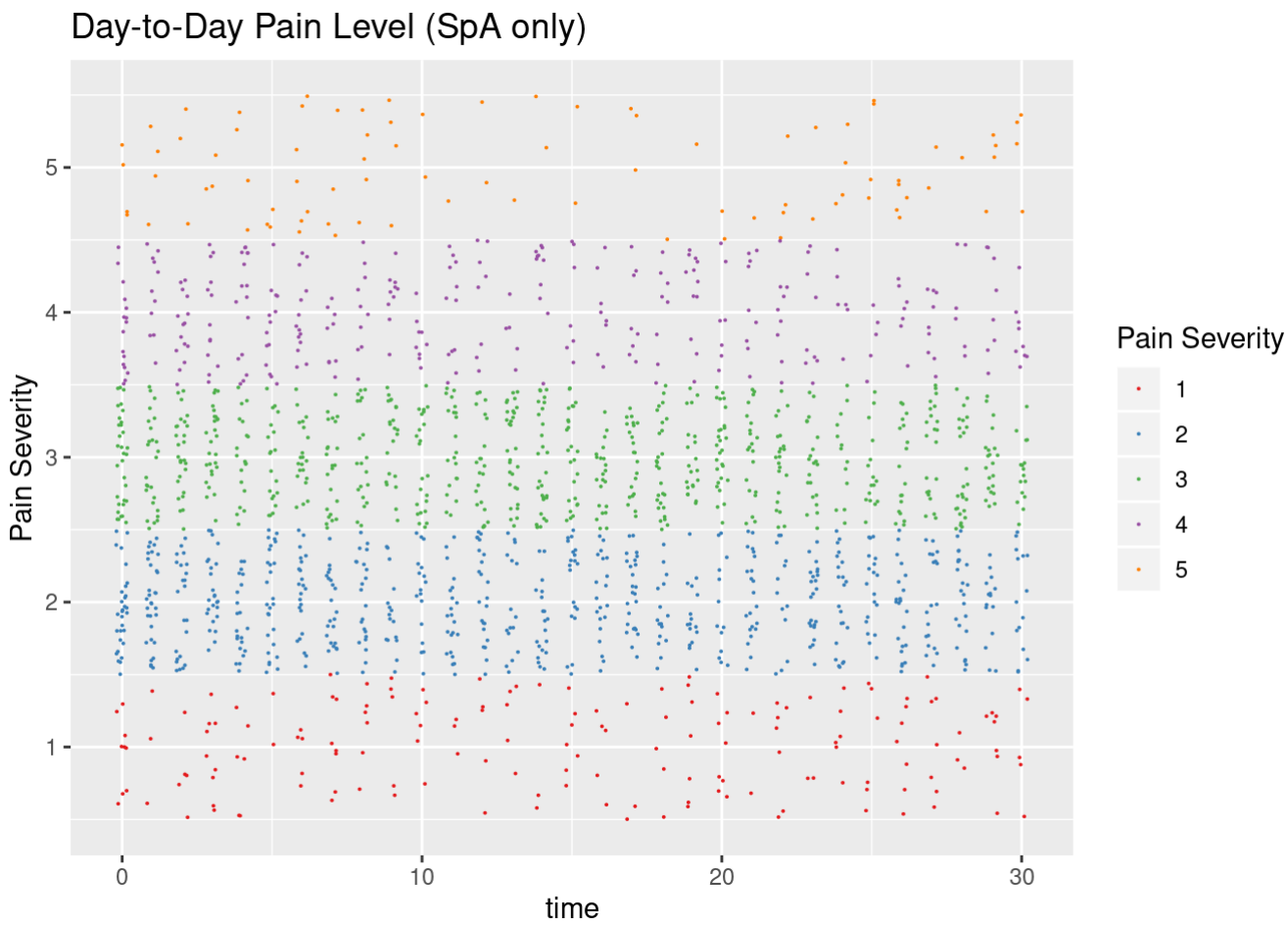
Day-to-day pain scoring for CWPFM (jitter plot)

Code Output



Day-to-day pain scoring for SpA (jitter plot)

Code Output



Markov Model (Observed State)

Pain severity category 0-5:

- 0 - missing data
- 1 - no pain
- 2 - mild pain
- 3 - moderate pain
- 4 - severe pain
- 5 - very severe pain

Transition matrix(missing data as a separate category)

Preparing the data

Code Output

```
##  [>] state coding:
```

```
##        [alphabet]  [label]  [long label]
```

```
##      1           0        0missing
```

```
##      2             1        1no pain
```

```
##      3             2        2mild pain
```

```
##      4             3        3moderate pain
```

```
##      5             4        4severe pain
```

```
##      6             5        5very severe pain
```

```
##  [>] 425 sequences in the data set
```

```
##  [>] min/max sequence length: 31/31
```

```
##  [>] state coding:
```

```
##        [alphabet]  [label]  [long label]
```

```
##      1             0        0missing
```

```
##      2             1        1no pain
```

```
##      3             2        2mild pain
```

```
##      4             3        3moderate pain
```

```
##      5             4        4severe pain
```

```
##      6             5        5very severe pain
```

```
##  [>] 409 sequences in the data set
```

```
##  [>] min/max sequence length: 31/31
```

```
##  [>] state coding:
```

```
##        [alphabet]  [label]  [long label]
```

```
##      1             0        0missing
```

```
##      2             1        1no pain
```

```
##      3              2        2mild pain
```

```
##      4              3        3moderate pain
```

```
##      5              4        4severe pain
```

```
##      6              5        5very severe pain
```

```
##  [>] 255 sequences in the data set
```

```
##  [>] min/max sequence length: 31/31
```

```
##  [>] state coding:
```

```
##        [alphabet]  [label]  [long label]
```

```
##      1              0        0missing
```

```
##      2              1        1no pain
```

```
##      3              2        2mild pain
```

```
##      4              3        3moderate pain
```

```
##      5              4        4severe pain
```

```
##      6              5        5very severe pain
```

```
##  [>] 100 sequences in the data set
```

```
##  [>] min/max sequence length: 31/31
```

Transition matrix 6x6 table for RA

Code Output

```
#### RA ####
mm_rawithNA <- build_mm(observations = ra_seqwithNA)

#### RA ####
mm_rawithNA$transition_probs %>%
  kable() %>%
  kable_styling(bootstrap_options = c("striped", "hover", "responsive", "condensed"))
```

Transition matrix 6x6 table for OA

Code Output

	0	1	2	3	4	5
0	0.5901956	0.0453997	0.1480319	0.1477904	0.0509539	0.0176286
1	0.2278177	0.3597122	0.2757794	0.0995204	0.0311751	0.0059952
2	0.2113665	0.0830140	0.4498723	0.2177522	0.0325670	0.0054278
3	0.2211337	0.0206438	0.2543737	0.3778866	0.1123163	0.0136459
4	0.2308438	0.0145490	0.1183317	0.3026188	0.2735209	0.0601358
5	0.2846715	0.0109489	0.0656934	0.1788321	0.2116788	0.2481752

Transition matrix 6x6 table for CWPFM

Code Output

	0	1	2	3	4	5
0	0.6665554	0.0186916	0.0904539	0.1188251	0.0771028	0.0283712
1	0.2307692	0.3576923	0.2538462	0.1153846	0.0230769	0.0192308
2	0.2476108	0.0660295	0.3640313	0.2649870	0.0451781	0.0121633
3	0.2271973	0.0143726	0.1763405	0.4002211	0.1580984	0.0237700
4	0.2286585	0.0050813	0.0640244	0.3099593	0.2916667	0.1006098
5	0.2222222	0.0066667	0.0288889	0.1088889	0.1977778	0.4355556

Transition matrix 6x6 table for SpA

Code Output

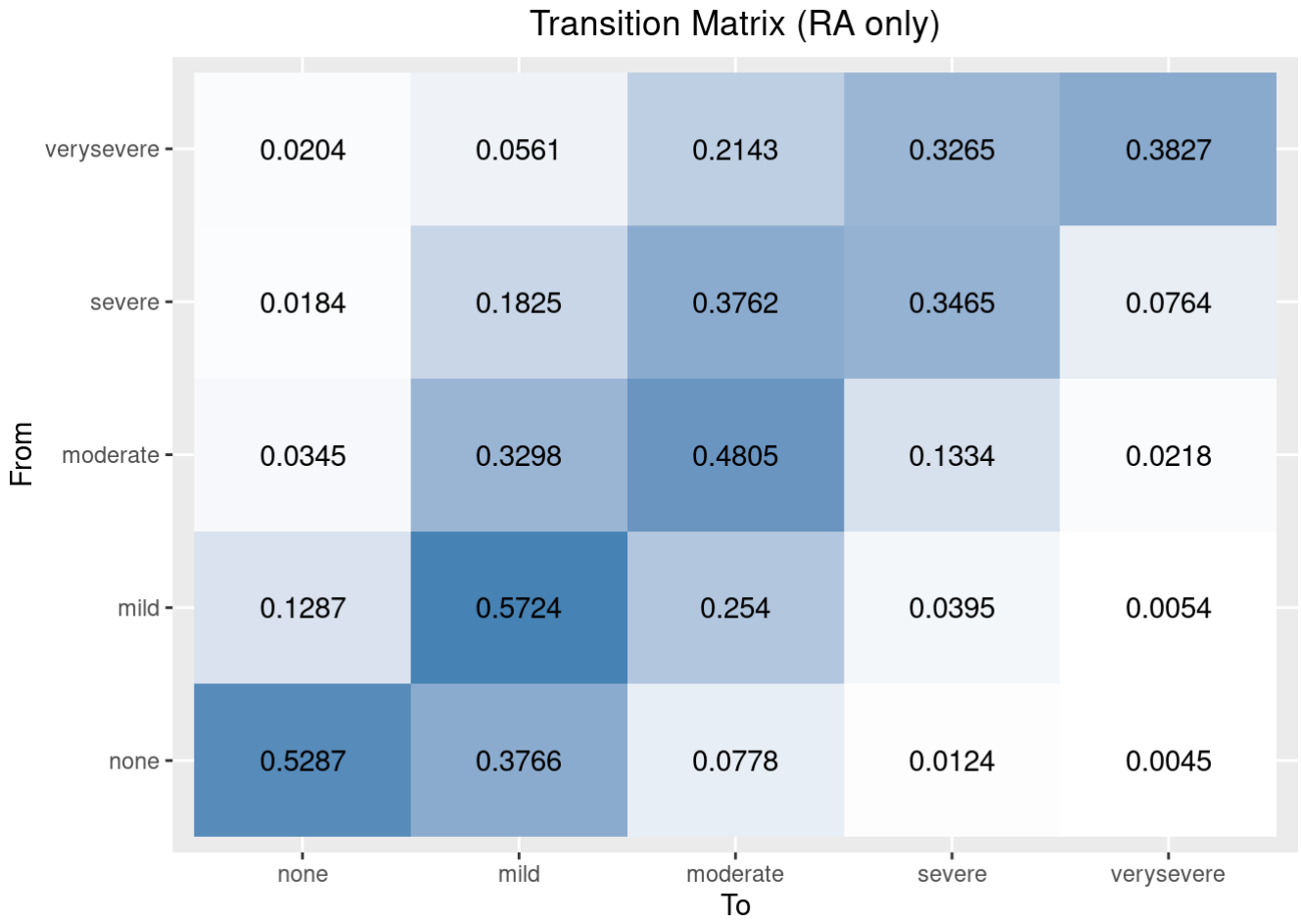
	0	1	2	3	4	5
0	0.6241259	0.0419580	0.1276224	0.1407343	0.0559441	0.0096154

	0	1	2	3	4	5
1	0.3190184	0.2576687	0.2638037	0.1104294	0.0368098	0.0122699
2	0.2639110	0.0747218	0.3926868	0.2193959	0.0413355	0.0079491
3	0.2558140	0.0174419	0.2063953	0.3779070	0.1264535	0.0159884
4	0.2268041	0.0206186	0.1202749	0.2852234	0.2714777	0.0756014
5	0.1411765	0.0470588	0.0117647	0.1411765	0.2588235	0.4000000

Transition probability heatmap plot for RA

- 1st plot: with missing category (6x6)
- 2nd plot: excluding missing category and re-calculation of the probability matrix (5x5)

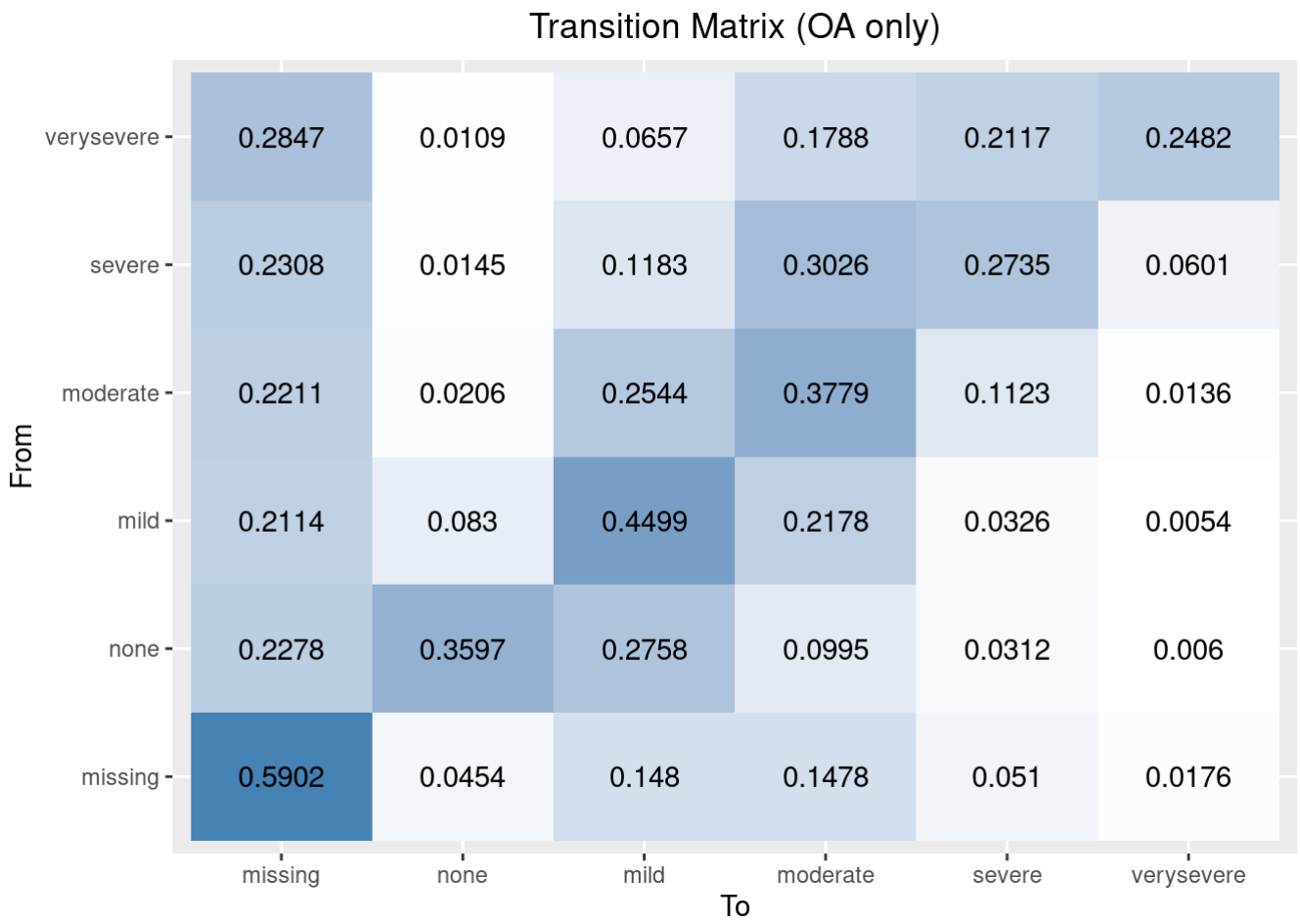
Code Output

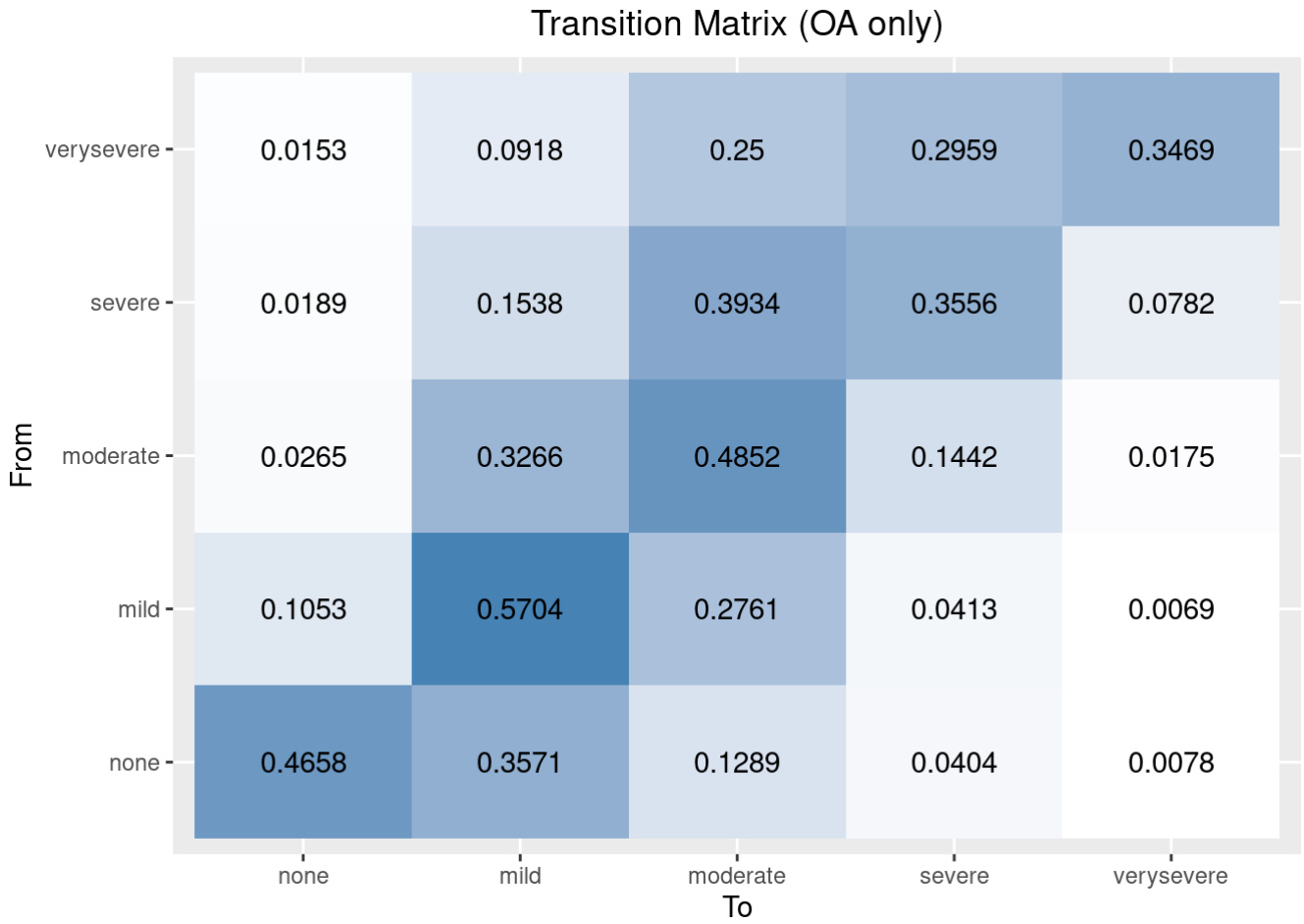


Transition probability heatmap plot for OA

- 1st plot: with missing category (6x6)
- 2nd plot: excluding missing category and re-calculation of the probability matrix (5x5)

Code Output

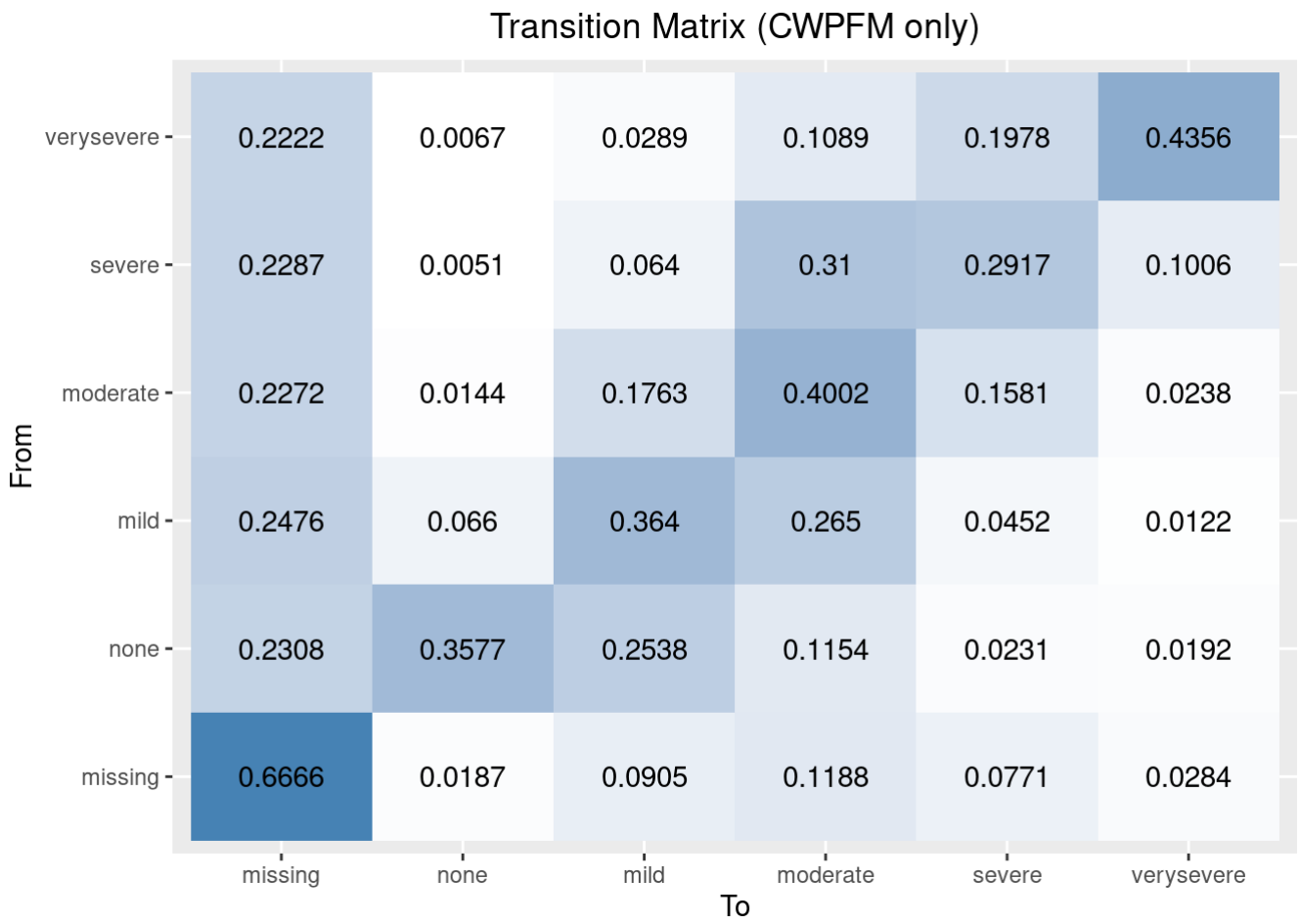

Transition Matrix (OA only)
From
verysevere
severe
moderate
mild
none
missing
0.2847 0.0109 0.0657 0.1788 0.2117 0.2482
0.2308 0.0145 0.1183 0.3026 0.2735 0.0601
0.2211 0.0206 0.2544 0.3779 0.1123 0.0136
0.2114 0.083 0.4499 0.2178 0.0326 0.0054
0.2278 0.3597 0.2758 0.0995 0.0312 0.006
0.5902 0.0454 0.148 0.1478 0.051 0.0176
missing
none
mild
moderate
severe
verysevere
To

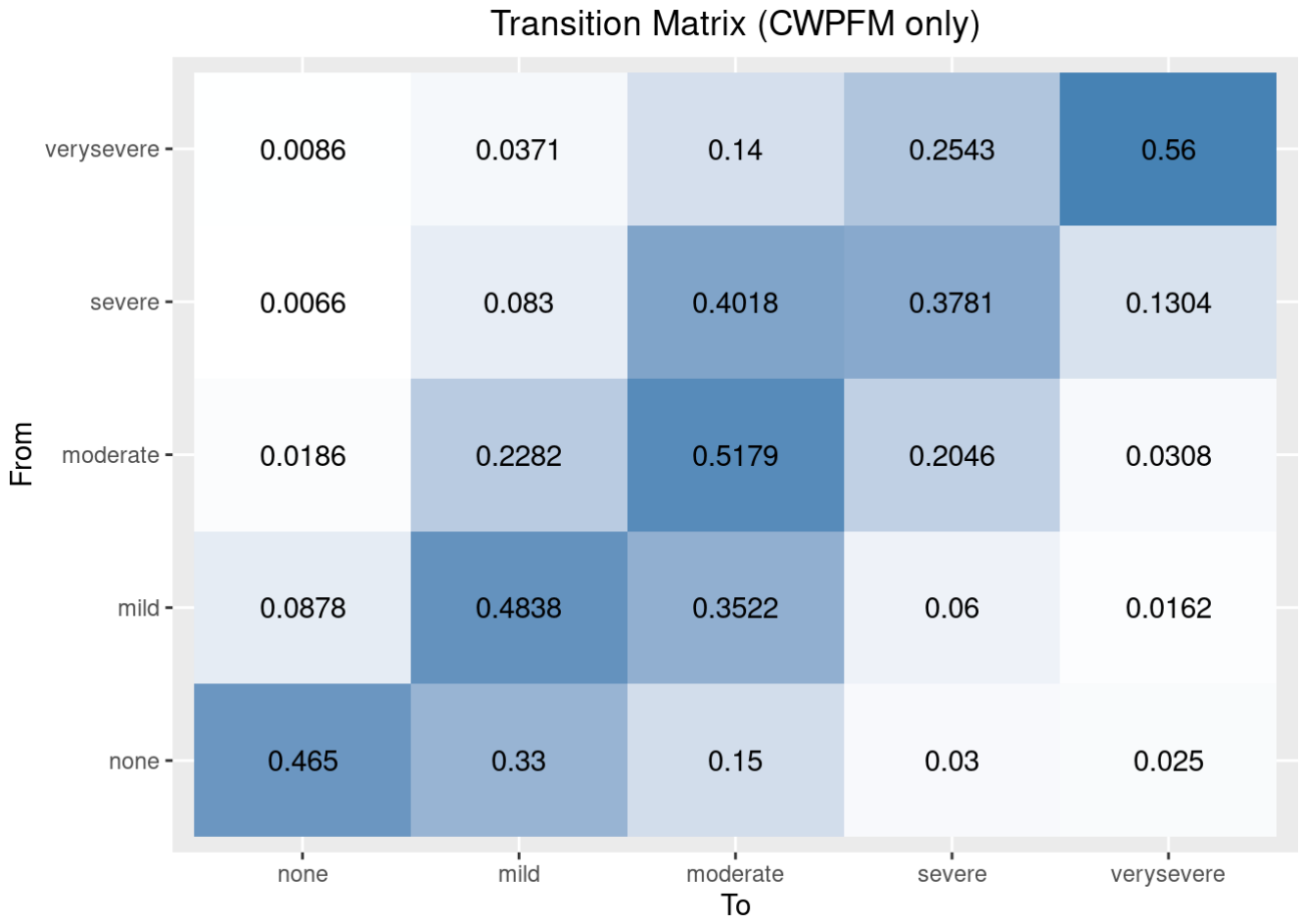

Transition Matrix (OA only)
From
verysevere
severe
moderate
mild
none
0.0153 0.0918 0.25 0.2959 0.3469
0.0189 0.1538 0.3934 0.3556 0.0782
0.0265 0.3266 0.4852 0.1442 0.0175
0.1053 0.5704 0.2761 0.0413 0.0069
0.4658 0.3571 0.1289 0.0404 0.0078
none
mild
moderate
severe
verysevere
To

Transition probability heatmap plot for CWPFM

- 1st plot: with missing category (6x6)
- 2nd plot: excluding missing category and re-calculation of the probability matrix (5x5)

Code Output

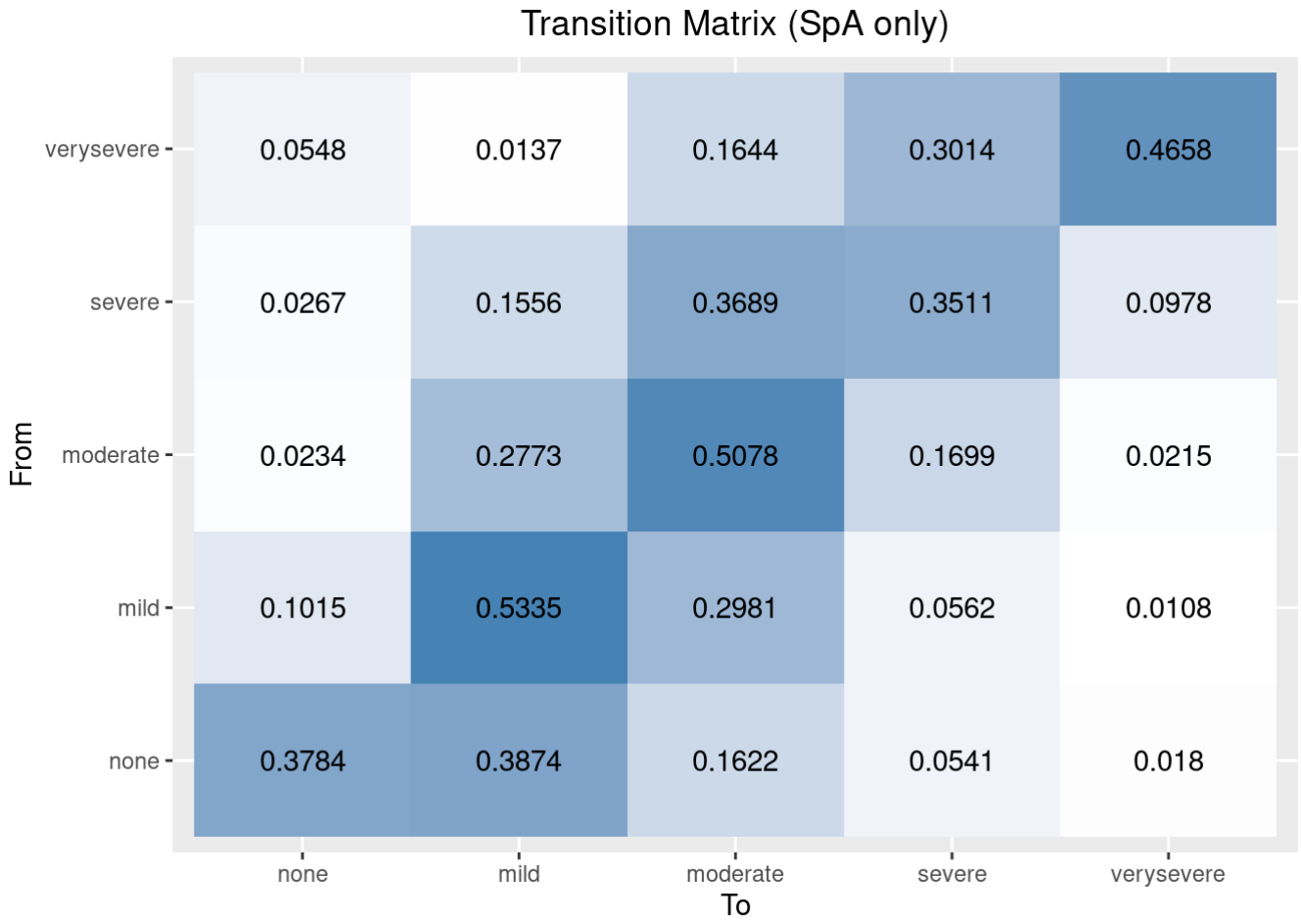




Transition probability heatmap plot for SpA

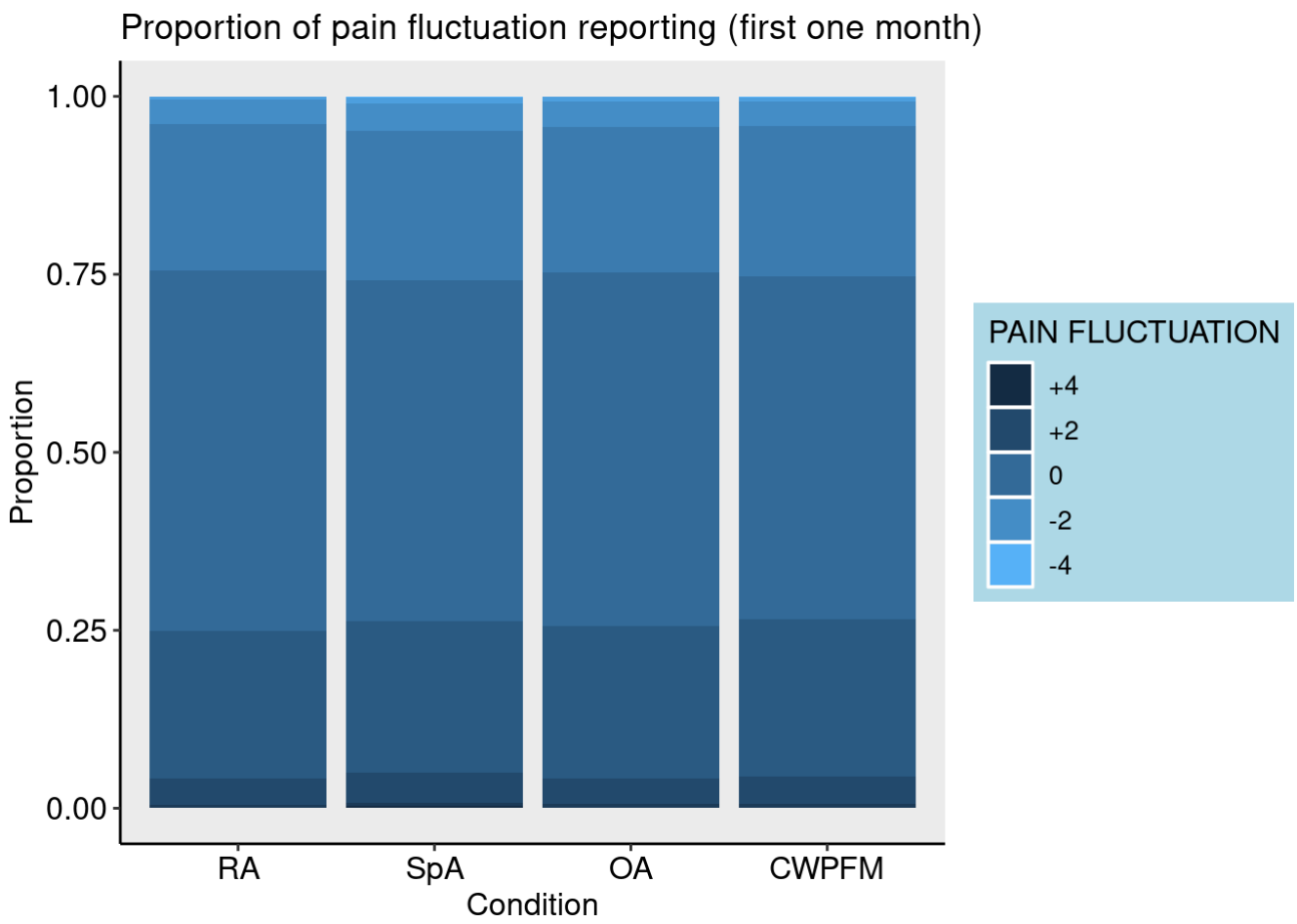
- 1st plot: with missing category (6x6)
- 2nd plot: excluding missing category and re-calculation of the probability matrix (5x5)

Code Output



PAIN FLUCTUATION

Code Output


Proportion of pain fluctuation reporting (first one month)
1.00
0.75
0.50
0.25
0.00
Proportion
RA
SpA
OA
CWPFM
Condition
PAIN FLUCTUATION
+4
+2
0
-2
-4

Appendix D: Three Primary Publications from *Cloudy with a Chance of Pain*

npj | Digital Medicine

www.nature.com/npjdigitalmed

ARTICLE OPEN

How the weather affects the pain of citizen scientists using a smartphone app

William G. Dixon[1,2,3*], Anna L. Beukenhorst[1], Belay B. Yimer[1], Louise Cook[1], Antonio Gasparrini[4,5], Tal El-Hay[6], Bruce Hellman[7], Ben James[7], Ana M. Vicedo-Cabrera[4], Malcolm Maclure[8], Ricardo Silva[9,10], John Ainsworth[2], Huai Leng Pisaniello[1,11], Thomas House[12,13], Mark Lunt[1], Carolyn Gamble[3,14,15], Caroline Sanders[14,15], David M. Schultz[16], Jamie C. Sergeant[1,3,17,18] and John McBeth[1,3,18]

Patients with chronic pain commonly believe their pain is related to the weather. Scientific evidence to support their beliefs is inconclusive, in part due to difficulties in getting a large dataset of patients frequently recording their pain symptoms during a variety of weather conditions. Smartphones allow the opportunity to collect data to overcome these difficulties. Our study *Cloudy with a Chance of Pain* analysed daily data from 2658 patients collected over a 15-month period. The analysis demonstrated significant yet modest relationships between pain and relative humidity, pressure and wind speed, with correlations remaining even when accounting for mood and physical activity. This research highlights how citizen-science experiments can collect large datasets on real-world populations to address long-standing health questions. These results will act as a starting point for a future system for patients to better manage their health through pain forecasts.

npj Digital Medicine (2019)2:105 ; https://doi.org/10.1038/s41746-019-0180-3

INTRODUCTION

Weather has been thought to affect symptoms in patients with chronic disease since the time of Hippocrates over 2000 years ago.[1] Around three-quarters of people living with arthritis believe their pain is affected by the weather.[2,3] Many report their pain is made worse by the cold, rain, and low atmospheric pressure. Others report that their pain is made worse by warmth and high humidity. Despite much research examining the existence and nature of the weather–pain relationship,[4] there remains no scientific consensus. Studies have failed to reach consensus in part due to their small sample sizes or short durations (commonly fewer than 100 participants or one month or less); by considering a limited range of weather conditions; and heterogeneity in study design (e.g. the populations studied, methods for assessing pain, assumptions to determine the weather exposure, and statistical analysis techniques).[5–11] Resolving this question requires collection of high-quality symptom and weather data on large numbers of individuals. Such data also need to include other factors potentially linked to daily pain variation and weather, such as mood and amount of physical activity. Collecting this kind of multi-faceted data in large populations over long periods of time, however, has been difficult.

The increasing uptake of smartphones offers new and significant opportunities for health research.[12] Smartphones allow the integration of data collection into daily life using applications (apps). Furthermore, embedded technologies within the smartphones, such as the Global Positioning System (GPS), can be used to link the data collection to specific locations. We created Cloudy with a Chance of Pain,[13,14] a national United Kingdom smartphone study, to collect a large dataset to examine the relationship between local weather and daily pain in people living with long-term pain conditions.

RESULTS

Recruitment and retention

The study app was downloaded by 13,207 users over the 12-month recruitment period (Figs 1 and 2a) with recruitment from all 124 UK postcode areas. A total of 10,584 participants had complete baseline information and at least one pain entry, with 6850 (65%) participants remaining in the study beyond their first week and 4692 (44%) beyond their first month (Fig. 2b). Further description of engagement clusters is provided in Supplementary Table 2 and Supplementary Figs 1–3. A total of 2658 participants had at least one hazard period matched to a control period in the same month (Fig. 3) and were included in the final analysis. There were 9695 hazard periods included in the analysis for the final 2658 participants, matched to 81,727 control periods in 6431 participant-months. A total of 1235 participants contributed one month, and the remaining 1423 participants contributed 2–15 months.

The final cohort was active for a median of 165 days (interquartile range, IQR 84–245) with symptoms submitted on an average of 73% of all days. Cohort members were

[1]Centre for Epidemiology Versus Arthritis, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK. [2]Health eResearch Centre, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK. [3]NIHR Greater Manchester Biomedical Research Centre, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK. [4]Department of Public Health Environments and Society, London School of Hygiene & Tropical Medicine, London, UK. [5]Centre for Statistical Methodology, London School of Hygiene & Tropical Medicine, London, UK. [6]IBM Research, Haifa, Israel. [7]uMotif Limited, London, UK. [8]Department of Anesthesiology, Pharmacology and Therapeutics, University of British Columbia, Vancouver, BC, Canada. [9]Department of Statistical Science, UCL, London, UK. [10]Alan Turing Institute, London, UK. [11]Discipline of Medicine, The University of Adelaide, Adelaide, Australia. [12]School of Mathematics, The University of Manchester, Manchester, UK. [13]IBM Research, Hartree Centre, Sci-Tech, Daresbury, UK. [14]Greater Manchester Patient Safety Translational Research Centre, The University of Manchester, Manchester, UK. [15]NIHR School of Primary Care, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK. [16]Centre for Atmospheric Science, School of Earth and Environmental Sciences, The University of Manchester, Manchester, UK. [17]Centre for Biostatistics, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK. [18]These authors contributed equally: Jamie C. Sergeant, John McBeth. *email: will.dixon@manchester.ac.uk

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

Fig. 1 User interface of the study app (uMotif, London). Each colored segment represents one of the ten data items. Participants report their symptoms on a five-point scale by dragging the segment from the center outwards

predominantly female (83%), had a mean age of 51 years (standard deviation 12.6), and had a range of different pain conditions, predominantly arthritis (Supplementary Table 1). The median number of weather stations associated with each participant during the course of their active data-collection period was 9 (IQR 4–14) with a maximum of 82 stations, indicating how mobile participants were during the course of the study and the importance of accounting for the weather at different locations over the course of the study. As an illustration of the structure of the data, the proportion of participants reporting a pain event was plotted as a heat map per calendar day for the study period (Fig. 4), aligned with the average United Kingdom weather data for the same time period. On any given day during the study, about 1–6% of participants had a pain event. At the start of the study, most participants believed in an association between weather and their pain (median score 8 out of 10, IQR 6–9). The demographics, health conditions and baseline beliefs of the 2658 participants included in the analysis were representative of the 10,584 participants who downloaded the app and provided baseline information (Supplementary Table 2).

Weather and pain

The multivariable case-crossover analysis including the four state weather variables demonstrated that an increase in relative humidity was associated with a higher odds of a pain event with an OR of 1.139 (95% confidence interval 1.099–1.181) per 10 percentage point increase, as was an increase in wind speed with an OR of 1.02 (1.005–1.035) per 1 m s^{-1} increase (Table 1). The odds of a pain event was lower with an increase in atmospheric pressure with an OR of 0.962 (0.937–0.987) per 10-mbar increase. Temperature did not have a significant association with pain (OR 0.996 (0.985–1.007) per 1 °C increase). The odds of a pain event was 12% higher per one standard deviation increase in relative humidity (9 percentage points) (OR 1.119 (1.084–1.154), compared to 4% lower for pressure (OR 0.958 (0.930–0.989) and 4% higher for wind speed (OR 1.041 (1.010–1.073) (11 mbar and 2 m s^{-1}, respectively). Of the four weather variables, relative humidity had the strongest association with pain, and temperature the least, evidenced by the estimated relative importance of the variables and their standardized odds ratios (Table 1, Supplementary Table 4). Similar effect sizes were seen when each variable was examined in univariable analyses. Precipitation was not associated with an increased odds of a pain event (OR 0.996 (0.989–1.003) per 1 mm daily rainfall amount) (Supplementary Table 5). Exploratory analyses considered time spent outside by including an interaction term with temperature, relative humidity, and wind speed. Time spent outside did not have a significant interaction with relative humidity or wind speed, nor did it lead to significant associations for temperature when conducting analyses stratified by time spent outside (Supplementary Table 3). It thus was not included in the final model.

The model was then expanded to include mood and physical activity on the day of interest, included as binary variables (Table 1), resulting in a modest reduction in the point estimates for all weather variables. Mood was strongly and independently associated with pain events (OR 4.083 (3.824–4.360) for low mood versus good mood), whereas there was no significant association with physical activity (OR 0.939 (0.881–1.002) for high versus low activity).

This multivariable regression model output represents the effect of one weather variable while all else remains constant. In reality, a single weather variable rarely changes in isolation while others remain unchanged. To illustrate the composite effect of the weather variables on the odds of reporting pain, an OR was calculated for each day using the coefficients of our multivariable model and daily UK mean weather values. Figure 5 demonstrates there is significant variability in the odds of a pain event for any given value of each weather variable. For example, at a temperature of 8 °C, the odds of a pain event varied from around

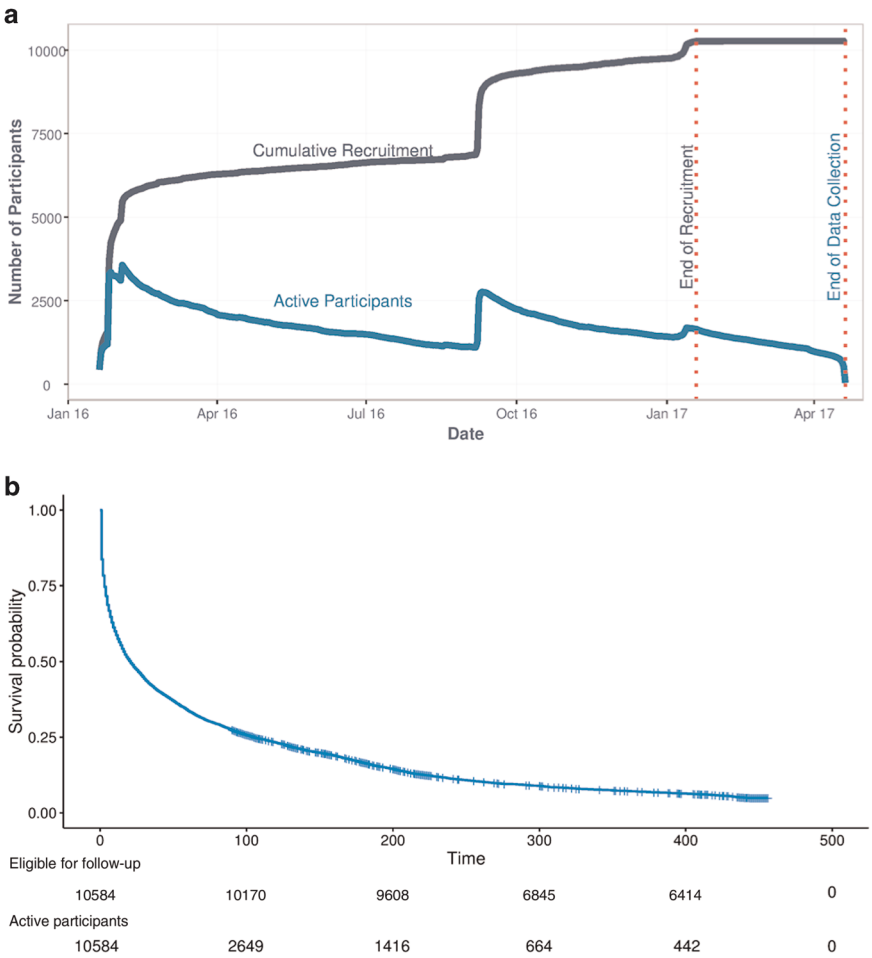


Fig. 2 Recruitment and retention. **a** Cumulative recruitment and number of active participants through time. The blue line represents the cumulative number of participants with a completed baseline questionnaire and at least one pain score submitted. The red line represents the current number of active participants (i.e. those who have submitted their first but not yet their last pain score in the study period). **b** Retention through time. The graph represents the retention of active participants through time as a survival probability from the day of their recruitment. Participants were censored when they were no longer eligible for follow-up. Eligible follow-up time ranged from 90 days (for those recruited on 20 January 2017) to 456 days (for those recruited on 20 January 2016)

0.8–1.2, depending on the other state variables in the weather that day.

Other factors such as day of the week (Supplementary Table 6), lagged weather values (Supplementary Table 7) and changes in weather variables from the previous day were tested. Mondays, Thursdays, and Saturdays (ORs 1.14, 1.14, and 1.29, respectively) had higher odds of pain compared to Sundays, but adjusting for the day of the week did not alter the effect of the four main weather variables. Except for relative humidity (1-day lag and 2-day lag), no significant associations were observed between lagged weather variables and pain events. Including change in weather from yesterday showed a minor effect of changing relative humidity (OR 1.005 (1.001–1.009) per 10 percentage point increase), whereas the effects of today's relative humidity and pressure remained unchanged (Supplementary Table 8). Stratification by disease led to a loss of statistical power and largely inconclusive results, although relative humidity appeared to have a stronger association with pain in patients with osteoarthritis (Supplementary Table 9, Supplementary Fig. 4). Stratification by the number of pain sites also showed no meaningful difference (Supplementary Table 10). After stratifying by participants' prior beliefs about their weather–pain relationship, relative humidity remained associated with pain in all participants although the association with pressure was only seen in those with a strong prior belief (Supplementary Table 11).

DISCUSSION

This study has demonstrated that higher relative humidity and wind speed, and lower atmospheric pressure, were associated with increased pain severity in people with long-term pain conditions. The most significant contribution was from relative humidity. The effect of weather on pain was not fully explained by its day-to-day effect on mood or physical activity. The overall effect sizes, while statistically significant, were modest. For example, the 'worst' combination of weather variables would increase the odds of a pain event by just over 20% compared to an average day. Nonetheless, such an increased risk may be meaningful to people living with chronic pain.

In addition to investigating the weather–pain relationship, we successfully conducted a national smartphone study that delivered on the promise of how consumer technology can support health research.[12,15] This study recruited over 10,000 participants throughout the United Kingdom, sustained daily self-reported data over many months,[13] and showcased the value of passively collected GPS data. Prior large smartphone studies have retained

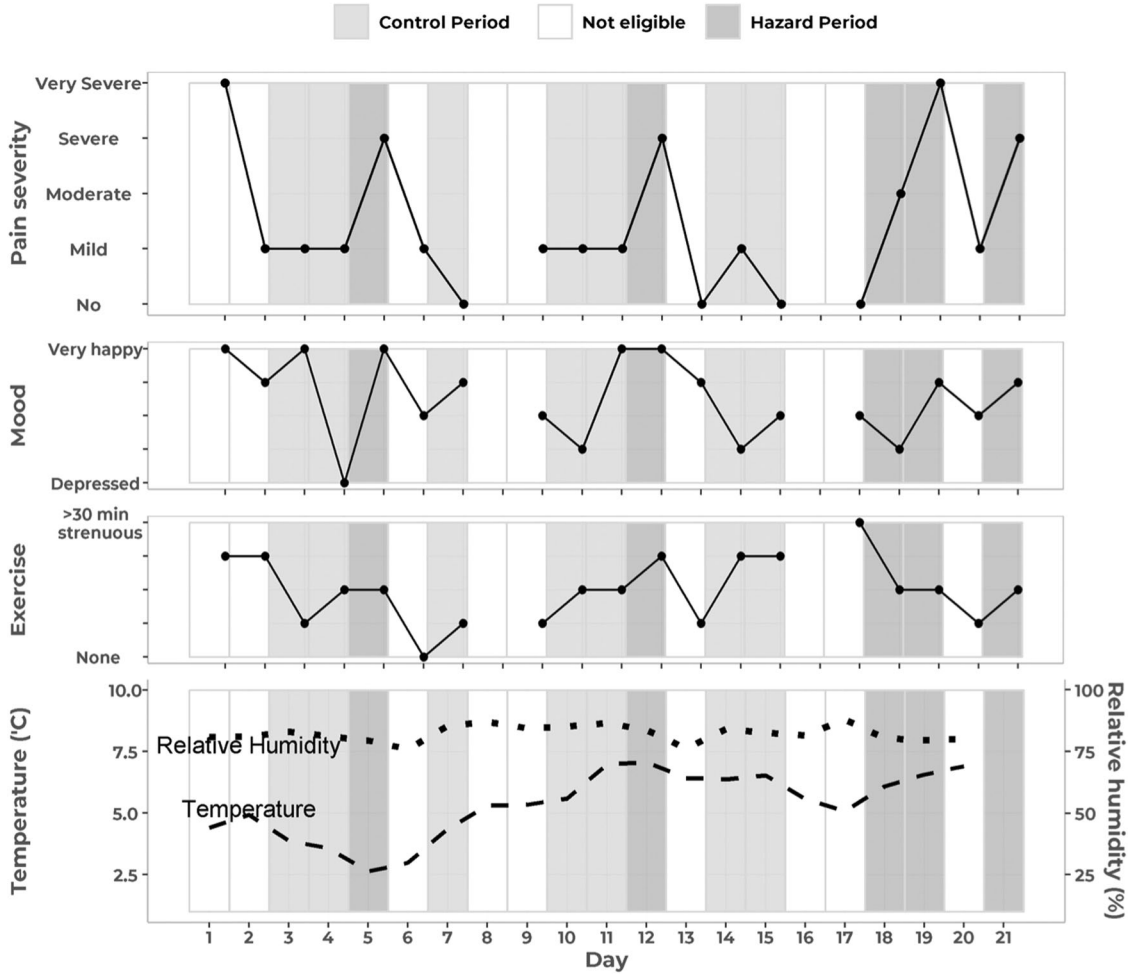


Fig. 3 Example participant timeline of 21 days, showing participant-reported items (here, pain severity, mood, and exercise) and weather data (here, temperature and relative humidity). Pain events with their associated hazard periods (dark grey) occur when pain severity increases by two or more ordinal categories between consecutive days (e.g. from Day 4 to Day 5). Control periods (light gray) occur on days that were eligible to be a pain event, but where pain did not increase by two or more ordinal categories. Days where there was no recorded pain on the preceding day, or where the preceding day's pain was severe or very severe (and could thus not increase by two or more categories), were not eligible to be pain-event days or control days. The case-crossover analysis compared the weather on pain-event days to weather on control days within a risk set of a calendar month

only around one in ten participants for seven days or less.[16,17] In contrast, our study retained 65% of participants for the first seven days, and 44% for the first month, with over 2600 participants contributing to the analysis having provided data for many months of the study.[13,14] An important success factor was strong public involvement in early setup and piloting, as well as participants' interest in weather as a possible pain trigger.[14] The study design has resolved problems of prior weather–pain studies such as small populations,[5,7] short follow-up,[3,8] surrogate pain outcomes,[11] the absence of possible causal pathway variables such as mood, and assumptions about where participants were located and thus the weather to which they were exposed.[18,19] Overcoming these obstacles produced a large dataset that allowed us to tease out subtle relationships between weather and pain.

There are potential limitations to this study. First, the reduction in participant numbers from over 10,000 with baseline data to the final 2658 participants with at least one within-month risk set raises questions about generalisability. Importantly, the characteristics of those included in the analysis were similar to the initial 10,000 participants, other than being slightly older (mean age 51 versus 48 years old). In a prior analysis, we showed that Cloudy participants were largely representative of a population reporting chronic-pain symptoms,[13] although proportionally fewer participants at both extremes of age were recruited. However, we would not expect middle-aged recruits to differ in their relationship between weather and pain from older or younger participants, and thus such selection factors would not invalidate our results. Second, the study was advertised to participants with a clear research question. It is possible that only people with a strong belief in a weather–pain relationship participated, generating an unrepresentative sample. However, the percentage of participants who believed in the weather–pain relationship was similar to prior studies,[20] and we did not see selective attrition of people who reported no weather–pain beliefs.[13] The within-person design would, regardless, mean that participants who drop out early would not introduce bias from time-invariant characteristics. Third, the lack of blinding raises possible information bias where observed weather could influence participants' symptom reporting. Our baseline questionnaire demonstrated that rain and cold weather were the most common pre-existing beliefs. If a reporting bias were to exist, we would expect higher pain to be reported at times of colder weather. Our findings—including the absence of an association with either temperature or rainfall—cannot be explained by such a reporting bias. Fourth, pain reporting is subjective, meaning one participant's "moderate" might equate to

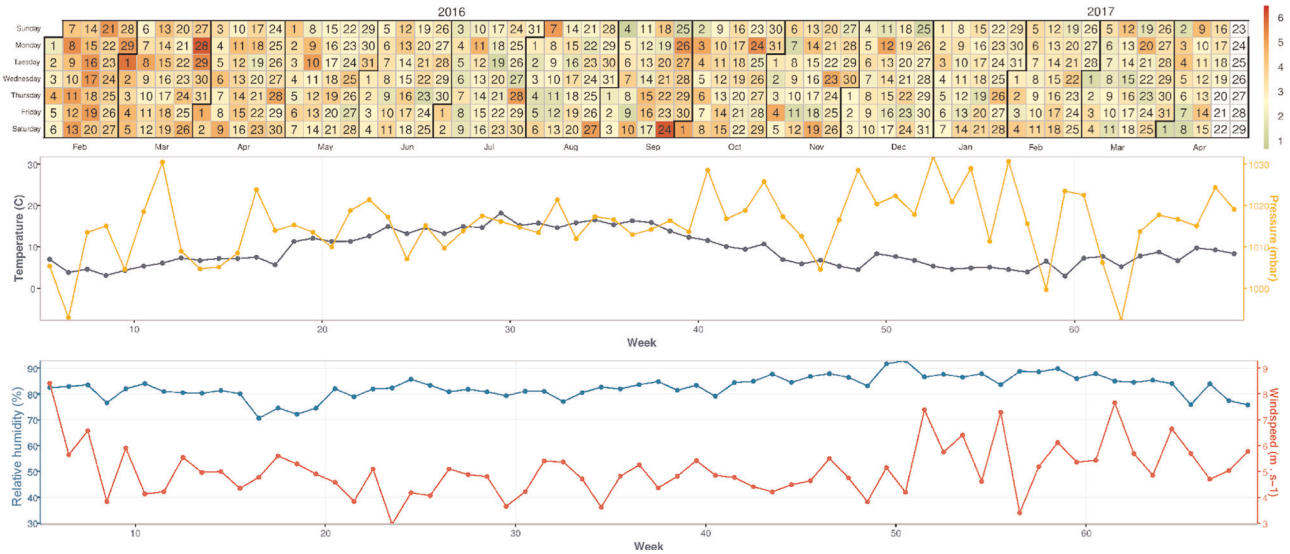


Fig. 4 The proportion of eligible active participants reporting a pain event during the study period, aligned with average UK weather data from February 2016 to April 2017. Heat map colors indicate the percentage of participants reporting a pain event on that day, ranging from 1–6% participants. The denominator per day is the number of participants who reported their pain on the day of interest and the prior day, irrespective of the level of pain on the prior day and thus their eligibility for a pain event

someone else's "severe". The within-person case-crossover analysis meant we compared moments when an individual's score increased by a meaningful amount to a control period for that same person. Fifth, we chose to model the weather using daily averages. It is possible that other findings may be hidden if the association between weather and pain was with other metrics of weather, such as the daily maximum, minimum, or range, or even if the changes in weather on hourly time scales affect participants' pain. Sixth, the findings from this United Kingdom study cannot necessarily be extrapolated to different climates where the weather is different. Seventh, our population-wide analysis assumed that all participants have the same weather–pain relationship. Different diseases may have different sensitivities to pain and, even within disease, participants may be affected differently. Our decision to use the whole chronic-pain population in our primary analysis means the overall associations with weather variables may be combinations of strong, weak and absent causal effects, thereby underestimating the most important associations. Notable differences were not seen after stratification by pain condition, although the power to detect any differences was reduced because of smaller sample sizes. Lastly, the inclusion of repeated events per person required us to consider within-subject dependence which, if not accounted for, would lead to bias.[21] Our outcome was based on changes in pain (a two or more category increase), which meant events rarely occurred on consecutive days, thereby ensuring a time gap between recurrent events and the avoidance of bias.

Understanding the relationship between weather and pain is important for several reasons. First, this study validates the perception of those who believe that their pain is associated with the weather. Second, given we can forecast the weather days in advance, understanding how weather relates to pain would allow pain forecasts. Patients could then plan activities and take greater control of their lives. Finally, understanding the relationship between weather and pain might also allow better understanding of the mechanisms for pain and thus allow the development of new and more effective interventions for those who suffer with pain.

In summary, our large national smartphone study has successfully supported the collection of daily symptoms and high-quality weather data, allowing examination of the relationship between weather and pain. The analysis has demonstrated significant relationships between relative humidity, pressure, wind speed and pain, with correlations remaining even when accounting for mood and physical activity.

METHODS

Patient involvement

Patient involvement has been important throughout the study, from inception to interpretation of the results. Co-author C.G. is a patient partner and co-applicant, while a patient and public involvement group of seven additional members has supported the study, meeting eight times in total. During the feasibility study,[14] patients positively influenced the wording and display of questions within the app. C.G. and other members of the Patient and Public Involvement Group were involved in media broadcasts at study launch and subsequent public engagement activities, explaining why the research question was important to them and relevant to patients with long-term pain conditions.[22] They have supported the interpretation of findings and the development of dissemination plans for the results, ensuring the results reach study participants, patient organizations and the general public.

Recruitment

We recruited participants through local and national media (television, radio, and press) and social media from 20 January 2016 to 20 January 2017. To participate in the study, participants needed to (i) be living with long-term (>3 months) pain conditions, (ii) be aged 17 years or older, (iii) be living in the United Kingdom, and (iv) own an Android or Apple iOS smartphone. Interested participants were directed to the study website (www.cloudywithachanceofpain.com) where they could check their eligibility, learn about the study, and download the uMotif app (Fig. 1). After downloading the study app, participants completed an electronic consent form and a baseline questionnaire including demographic information (sex, year of birth, first half of postcode), anatomical site(s) of pain, underlying pain condition(s), baseline medication use, and beliefs about the extent to which weather influenced their pain on a scale of 0–10, including which weather condition(s) were thought to be most associated with pain. Participants were then invited to collect daily symptoms for six months, or longer if willing. Each day, the app alerted participants to complete ten items at 6:24 p.m. (Fig. 1). The ten items were pain severity, fatigue, morning stiffness, impact of pain, sleep quality, time spent outside, waking up feeling tired, physical activity, mood, and well-being. Each data

Table 1. Association between weather and pain from the case-crossover analysis in 2658 participants

Variable	Univariable (single weather variable only) Odds ratio (95% CI)	Multivariable (all weather variables only) Odds ratio (95% CI)	Multivariable (weather plus activity and mood) Odds ratio (95% CI)
Temperature			
Per 1 °C	1.001 (0.991–1.012)	0.996 (0.985–1.007)	1.001 (0.989–1.013)
Per 1 s.d. (4.8 °C)	1.007 (0.956–1.060)	0.981 (0.929–1.035)	1.005 (0.949–1.064)
Relative humidity			
Per 10%	1.148 (1.108–1.189)	1.139 (1.099–1.181)	1.117 (1.075-1.16)
Per 1 s.d. (8.6%)	1.126 (1.092–1.61)	1.119 (1.084–1.154)	1.100 (1.064–1.136)
Pressure			
Per 10 mbar	0.936 (0.914–0.958)	0.962 (0.937–0.987)	0.966 (0.94-0.993)
Per 1 s.d. (11.1 mbar)	0.930 (0.905–0.955)	0.958 (0.930–0.986)	0.963 (0.934–0.992)
Wind speed			
Per 1 m s^{-1}	1.023 (1.01–1.037)	1.02 (1.005–1.035)	1.011 (0.995–1.027)
Per 1 s.d. (2.1 m s^{-1})	1.048 (1.020–1.077)	1.041 (1.010–1.073)	1.022 (0.990–1.056)
High activity			0.939 (0.881–1.002)
Low mood			4.083 (3.824–4.360)

High activity—Top three categories: 30 min or more of light or strenuous activity per day, or less than 30 min of strenuous activity
Low mood—Bottom three categories: 'depressed', 'feeling low' or 'not very happy' *s.d.* standard deviation
Distribution of weather variables:
Temperature: range −4.9 to 25.9 °C, s.d. 4.8 °C
Relative humidity: range 43.8–100%, s.d. 8.6%
Pressure: range 966–1044.8 mbar, s.d. 11.1 mbar
Wind speed: range 0–21.5 m s^{-1}, s.d. 2.1 m s^{-1}

item had five possible labeled ordinal responses. For example, in response to the question "How severe was your pain today?", possible responses were "no pain", "mild pain", "moderate pain", "severe pain" or "very severe pain". The data were analysed using a case-crossover design where, for each participant, exposure during days with a pain event ("hazard periods") were compared to "control periods" without a pain event in the same month.[23] Pain events were defined as a two-or-more category increase in pain from the preceding day, consistent with more stringent definitions of a clinically important difference[24] (Fig. 3). Data collection ended on 20 April 2017.

Cohort selection

Participants were included in the final cohort for analysis if they fulfilled the following criteria: (1) downloaded the app; (2) provided consent; (3) completed the baseline questionnaire; and (4) contributed at least one pain event and matched control period in the same month (see below). During exploratory analysis, it was apparent that people reported higher pain levels in the first ten days following recruitment (perhaps due to calibration or regression to the mean). Therefore, the first ten days were excluded from the formal analysis. However, even if the first ten days were included, they had a negligible effect on the results (Supplementary Table 12).

The total person-days in study was calculated for each participant as the number of days between their first and last day of entering pain data. The number of person-days on which a pain score was entered was summed per participant, presented as a proportion of the total person-days in study, and averaged across the population. The geographical distribution of recruitment was described as the number of UK postcode areas represented (out of a maximum of 124).[25] The movement of participants during the study was described as the median number of weather stations associated with each participant during their data-collection period.

Ethical approval

Ethical approval was obtained from the University of Manchester Research Ethics Committee (ref: ethics/15522) and from the NHS IRAS (ref: 23/NW/0716). Participants were required to provide electronic consent for study inclusion. Further details are available elsewhere.[13,14]

Weather data

Weather data were obtained by linking hourly smartphone GPS data to the nearest of 154 possible United Kingdom Met Office weather stations. Where GPS data were missing, we used significant location imputation. (For details, see supplement). Local hourly weather data were obtained from the Integrated Surface Database (ISD) of NOAA (http://www.ncdc.noaa.gov/isd), which includes hourly observations from UK Met Office weather stations.

Given the latitude–longitude coordinates of a participant location, the haversine distance to every Met Office weather station was calculated. The nearest station to the given location was selected, conditional on the distance being less than 100 km and the station having four weather variables (temperature, pressure, wind speed, and dewpoint temperature) available at that time. If all stations with the required weather data exceeded the maximum distance (100 km), the location was left unlinked and the observation was excluded from the analysis.

The significant location imputation approach for handling missing hourly GPS data had three stages.[26] First, the participant's observed location data were ordered by the frequency that the locations were visited. Second, the locations were spatially clustered using Hartigan's Leader Algorithm[27] with a threshold of 0.5 km. Third, missing locations during weekdays were replaced by the centroid of the participant's most visited cluster for weekdays and missing locations during weekends were replaced by centroid of the participant's most visited cluster for weekends.

Recruitment and retention

Recruitment and duration of follow-up were presented as a graph of cumulative recruitment and active participants, with participation ending at the last symptom entry. Retention in the study was also presented as a survival probability against time since recruitment, with participants censored when they were no longer eligible for follow-up. Eligible follow-up time ranged from 90 days (for those recruited on 20 January 2017) to 456 days (for those recruited on 20 January 2016). Engagement of

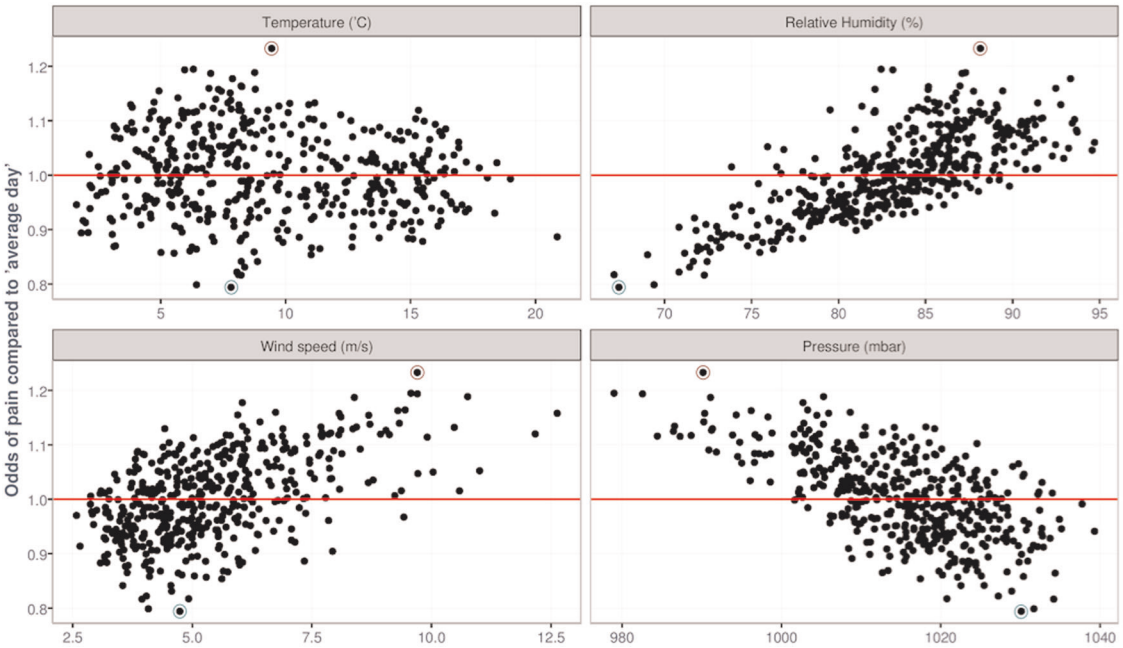


Fig. 5 Estimated odds of a painful day for all weather days experienced during the 15 months. Estimated odds of a painful day are plotted as the odds ratio for each day compared to the average weather day in this period (temperature = 9.3 °C, relative humidity = 83%, wind speed = 4 m s^{-1} and pressure = 1013 mbar). Estimated odds are calculated from the output of the multivariable regression analysis. The day associated with the highest estimated odds of a pain event had a temperature of 9 °C, relative humidity 88%, wind speed 9.5 m s^{-1} and pressure 988 mbar. The day associated with the lowest estimated odds of a pain event was when the temperature was 7 °C, relative humidity was 67%, wind speed 4.5 m s^{-1} and pressure 1030 mbar

participants was further described through clustering of engagement states, which has been described in detail elsewhere.[13] Following recruitment, individuals were labeled as engaged if they reported any of the ten symptoms on a given day. A first-order hidden Markov model was used to estimate the levels of engagement of participants by assuming three latent engagement states: high, low, and disengaged. Clusters were defined according to different probabilities of transitioning between high engagement, low engagement and disengagement during the study. Retention of active participants was also presented stratified by engagement cluster, and in the subset of participants who contributed to the final analysis.

Analysis method

Days without pain events were only control periods if they were eligible to have a two-or-more category increase (i.e. the preceding day's pain was lower than "severe"), thus fulfilling the exchangeability assumption for the case-crossover study design.[28] With this design, participants serve as their own control, eliminating confounding by time-invariant factors. Each month per participant with at least one hazard and one control period formed a risk set. Conditional logistic regression was used to estimate the odds ratio (OR) for a pain event for four state weather variables (temperature, relative humidity, pressure, and wind speed). The condition logistic regression model was implemented with the assumption that the possible recurrent events (hazard periods) within a person are independent conditional on the subject-specific variables and other observed time-varying covariates. Further, we make sure that there is no overlap between case and control periods. Our assumption is reasonable given the time gap between subsequent events.

Each weather variable was included in univariable models and then all four were included in a multivariable analysis. Each weather variable was represented as a daily average per participant for the hazard or control period, with results presented as an OR for a pain event in response to a one-unit increase for temperature and wind speed (°C and meter per second, respectively) or a ten-unit increase for relative humidity and pressure (percentage points and millibar, respectively). Standardized odds ratios of each weather variable were also calculated. The relative importance of the four state weather variables was estimated by summing the Akaike weights.[29] In all models, the preceding day's pain score was included as it influenced the likelihood of a pain event the following day. The model was expanded to include mood and physical activity on the day of interest, included as binary variables. Time spent outside was considered as a possible effect modifier by including an interaction term with temperature, relative humidity, and wind speed. A directed acyclic graph is included in the supplementary material (Supplementary Fig. 5).

Sensitivity analyses were conducted to examine the effect of precipitation, day of the week, possible lag between weather and pain, change in weather from the day before the hazard or control day, disease type, sites of pain (single versus multiple sites) and prior beliefs in the weather–pain relationship. Respecting patients' perspectives, we decided our primary analysis would focus on the whole chronic-pain population and our analyses of disease-specific associations would be secondary. We also re-ran the analysis including the first 10 days.

Daily pain-event estimates

Estimated odds ratio for a pain event per day compared to the average weather day were calculated using the following equation:

$$\text{Odds Ratio} = \exp,\ \Big\{\beta_T(\text{temperature} - \mu_T) + \beta_{RH}\left(\frac{\text{relative humidity} - \mu_{RH}}{10}\right) + \beta_{wsp}(\text{wind speed} - \mu_{wsp}) + \beta_P\left(\frac{\text{pressure} - \mu_P}{10}\right)\Big\}$$

where
β_T = coefficient for temperature from final model,
β_{RH} = coefficient for relative humidity,
β_{wsp} = coefficient for wind speed,
β_P = coefficient for pressure, and
μ_T = mean temperature,
μ_{RH} = mean relative humidity,
μ_{wsp} = mean wind speed, and
μ_p = mean pressure
of the daily UK means over the study period.

The predicted odds ratios of a pain event, relative to the average weather day, were plotted for all days within our study period for each of the four state weather variables.

Statistical analyses were performed using R 3.3.0.[30]

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

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 3.3.0. Code may be available on reasonable request.

Received: 22 May 2019; Accepted: 23 September 2019;
Published online: 24 October 2019

REFERENCES

1. Hippocrates. On Airs, Waters and Places. http://classics.mit.edu//Hippocrates/airwatpl.html. (400 B.C.E.).
2. Hagglund, K. J. et al. Weather, beliefs about weather, and disease severity among patients with fibromyalgia. *Arthritis Rheum.* **7**, 130–135 (1994).
3. Timmermans, E. J. et al. Self-perceived weather sensitivity and joint pain in older people with osteoarthritis in six European countries: Results from the European Project on OSteoArthritis (EPOSA). *BMC Musculoskelet. Disord.*, https://doi.org/10.1186/1471-2474-15-66 (2014).
4. Smedslund, G. & Hagen, K. B. Does rain really cause pain? A systematic review of the associations between weather factors and severity of pain in people with rheumatoid arthritis. *Eur. J. Pain.* **15**, 5–10 (2011).
5. Aikman, H. The association between arthritis and the weather. *Int. J. Biometeorol.* **40**, 192–199 (1997).
6. Brennan, S. A. et al. Influence of weather variables on pain severity in end-stage osteoarthritis. *Int. Orthop.* **36**, 643–646 (2012).
7. Smedslund, G. et al. Does the weather really matter? A cohort study of influences of weather and solar conditions on daily variations of joint pain in patients with rheumatoid arthritis. *Arthritis Rheum.* **61**, 1243–1247 (2009).
8. Bossema, E. R. et al. Influence of weather on daily symptoms of pain and fatigue in female patients with fibromyalgia: a multilevel regression analysis. *Arthritis Care Res. (Hoboken)* **65**, 1019–1025 (2013).
9. Duong, V. et al. Does weather affect daily pain intensity levels in patients with acute low back pain? A prospective cohort study. *Rheumatol. Int.* **36**, 679–684 (2016).
10. Guedj, D. & Weinberger, A. Effect of weather conditions on rheumatic patients. *Ann. Rheum. Dis.* **49**, 158–159 (1990).
11. Jena, A. B. et al. Association between rainfall and diagnoses of joint or back pain: retrospective claims analysis. *BMJ* **359**, j5326 (2017).
12. Hayden, E. C. Mobile-phone health apps deliver data bounty. *Nature* **531**, 422–423 (2016).
13. Druce, K. L. et al. Recruitment and ongoing engagement in a UK smartphone study examining the association between weather and pain: Cohort study. *JMIR mHealth uHealth* **5**, e168 (2017).
14. Reade, S. et al. Cloudy with a Chance of Pain: Engagement and subsequent attrition of daily data entry in a smartphone pilot study tracking weather, disease severity, and physical activity in patients with rheumatoid arthritis. *JMIR mHealth uHealth* **5**, e37 (2017).
15. Santos-Lozano, A. et al. mHealth and the legacy of John Snow. *Lancet* **391**, 1479–1480 (2018).
16. Bot, B. M. et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. *Sci. Data* **3**, 160011 (2016).
17. McConnell, M. V. et al. Feasibility of obtaining measures of lifestyle from a smartphone app. *JAMA Cardiol.* **2**, 67 (2017).
18. Smedslund, G. et al. Do weather changes influence pain levels in women with fibromyalgia, and can psychosocial variables moderate these influences? *Int. J. Biometeorol.* **58**, 1451–1457 (2014).
19. Beilken, K. et al. Acute low back pain? Do not blame the weather—A case-crossover study. *Pain. Med.* **18**, pnw126 (2016).
20. Jamison, R. N., Anderson, K. O. & Slater, M. A. Weather changes and pain: perceived influence of local climate on pain complaint in chronic pain patients. *Pain* **61**, 309–315 (1995).
21. Luo, X. & Sorock, G. S. Analysis of recurrent event data under the case-crossover design with applications to elderly falls. *Stat. Med.* **27**, 2890–2901 (2008).
22. Cloudy with a Chance of Pain on BBC North West Tonight. https://www.youtube.com/watch?v=YUdtKGr49GY. Accessed 14 Oct 2019 (2016).
23. Maclure, M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am. J. Epidemiol.* **133**, 144–153 (1991).
24. Olsen, M. F. et al. Pain relief that matters to patients: systematic review of empirical studies assessing the minimum clinically important difference in acute pain. *BMC Med.* **15**, 35. https://doi.org/10.1186/s12916-016-0775-3 (2017).
25. BPH postcodes. A brief guide to UK postcodes. https://www.bph-postcodes.co.uk/guidetopc.cgi (2018).
26. Isaacman, S., Becker, R., Martonosi, M., Rowland, J., & Varshavsky, A. Identifying important places in people's lives from cellular network data sibren. *Proc. of 9th Int. Conf. Pervasive Comput.* 1–18 (2011).
27. Hartigan, J. A. *Clustering Algorithms*. (Wiley, New York, 1975).
28. Mittleman, M. A. & Mostofsky, E. Exchangeability in the case-crossover design. *Int. J. Epidemiol.* **43**, 1645–1655 (2014).
29. Burnham, K. P., Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd ed. (Springer-Verlag, New York, 2003).
30. R Core Team. R: A language and environment for statistical computing, https://www.r-project.org (2018).

ACKNOWLEDGEMENTS

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. The study app and website was provided by uMotif Limited (London, UK). The unique flower-like 'motif' symptom tracking interface is owned by uMotif Limited and protected through EU Design Registrations and a U.S. Design Patent. We gratefully acknowledge the National Oceanic and Atmospheric Administration/National Climatic Data Center Integrated Surface Database (https://www.ncdc.noaa.gov/isd) for providing the weather data used in this study. The study was funded by Versus Arthritis (new name for Arthritis Research UK) (grant reference 21225), with additional support from the Centre for Epidemiology (grants 21755 and 20380). A.G. and A.M.V.C. are the recipients of Medical Research Council U.K. grants (MR/M022625/1 and MR/R013349/1). H.L.P. is the recipient of the Ken Muirden Overseas Training Fellowship from the Arthritis Australia, an educational research grant funded by the Australian Rheumatology Association. A.B. is supported by a Medical Research Council doctoral training partnership (grant MR/N013751/1). T.H. is supported by the Alan Turing Institute and the Royal Society (grant INF/R2/180067). D.M.S. is partially supported by the Natural Environment Research Council U.K. (grants NE/I005234/1, NE/I026545/1, and NE/N003918/1). R.S. is partially supported by the Alan Turing Institute (grant EP/N510129/1).

AUTHOR CONTRIBUTIONS

W.G.D. designed the study, acquired funding, supervised and participated in data-collection and content analysis, and wrote the first draft of the manuscript. A.L.B., B.B.Y. and H.L.P. conducted the analysis. L.C. coordinated project management and participant support. A.G., T.E.L., A.V.M.C., M.M., R.S., T.H., M.L., D.M.S., J.C.S. and J. McB. contributed to analysis plans and supervised the analysis. B.H., B.J., J.A., C.G., C.S., D.M.S., J.C.S. and J.McB. contributed to study design. C.S. led qualitative research in the feasibility study and led patient and public involvement. All authors critically reviewed manuscript drafts and approved the final version of the manuscript. W.G.D. is responsible for the overall content as guarantor, and attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

COMPETING INTERESTS

W.G.D. has received consultancy fees from Bayer Pharmaceuticals and Google, unrelated to this study. B.J. and B.H. are co-founders of uMotif. All other authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41746-019-0180-3.

Correspondence and requests for materials should be addressed to W.G.D.

Reprints and permission information is available at http://www.nature.com/reprints

Cloudy with a chance of Pain

Adapted from "Weather Patterns Associated with Pain in Chronic-Pain Sufferers," by **David M. Schultz** (University of Manchester), **Anna L. Beukenhorst**, **Belay Birlie Yimer**, **Louise Cook**, **Huai Leng Pisaniello**, **Thomas House**, **Carolyn Gamble**, **Jamie C. Sergeant**, **John McBeth**, and **William G. Dixon**. Published online in *BAMS*, May 2020. For the full, citable article, see DOI:10.1175/BAMS-D-19-0265.1.

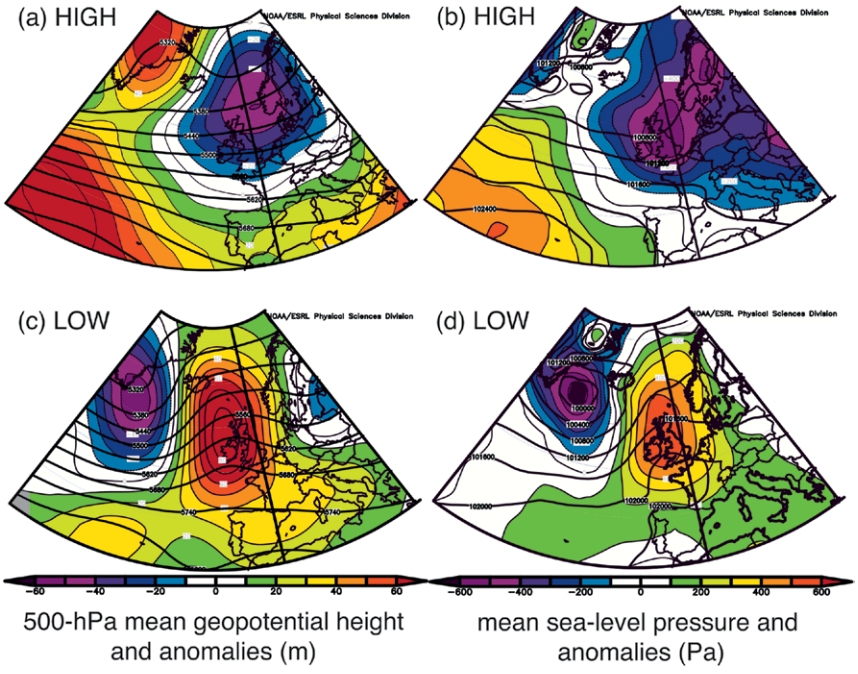
A common belief among three-quarters of those who suffer from chronic pain is that their pain fluctuates with the weather. However, this belief lacks scientific support. In a recent review of 41 studies examining the relationship between weather and chronic musculoskeletal pain, 28 (68%) found some relationship, although there was disagreement about what that relationship was.

Various reasons have been postulated for the disagreement: small sample sizes as measured by the number of participants, the duration of the study, or both; weather observations that were unrepresentative of conditions experienced by participants; and the lack of input or analysis by meteorologists.

To overcome these limitations of previous studies, we designed a 15-month-long U.K.-based citizen-science smartphone project called Cloudy with a Chance of Pain (www.cloudywithachanceofpain.com). Participants with chronic pain used a specially designed app to enter a 10-question daily report on their pain and other well-being characteristics on a 5-point scale. For example, participants rated their pain from 1 ("no pain") to 5 ("very severe pain").

The absolute number in a participant's report may not be meaningful, by itself: participants reporting the same severity on our 5-point scale may experience different levels of pain. Studies show, however, that a 20% *increase* in pain severity is clinically significant. As such, we defined a pain event for an individual participant when they report a 1-category or greater increase (+1) in their pain level from the previous day.

Using the global positioning system (GPS) sensor in the phone, we linked the participants' locations during the study to the closest weather stations in the Met Office observing network. Thus, we developed a daily profile of the average weather conditions each



Synoptic composites of HIGH and LOW days. Anomalies computed relative to the weighted average of the daily 1981–2010 means.

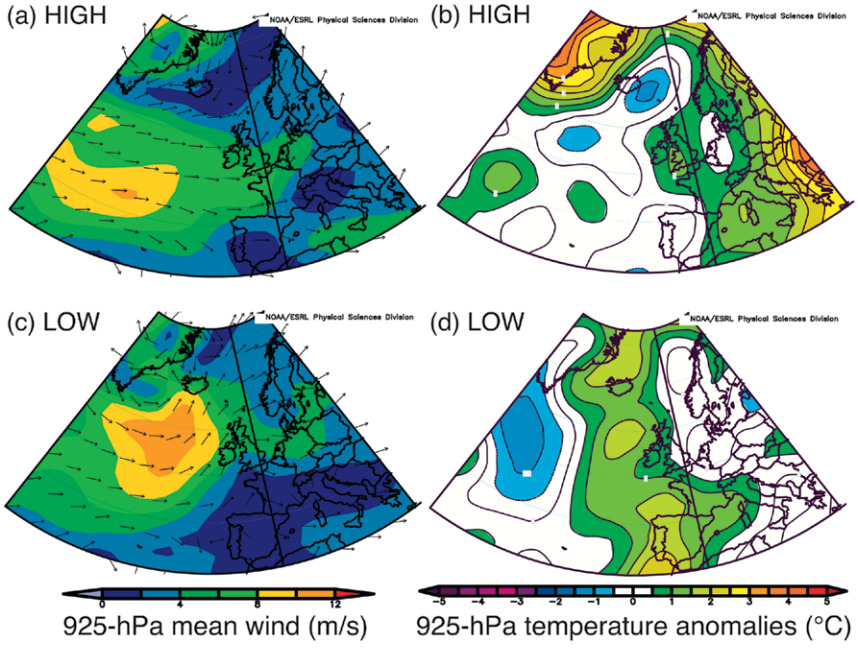
Data and methods

The study ran from 20 January 2016 to 19 April 2017. Bolstered by national media attention, 13,207 users throughout the United Kingdom downloaded the study app at some point during a 12-month recruitment period. A total of 10,584 participants entered their demographic information and at least one pain report, making them eligible for the study. Sixty-five percent of participants remained in the study beyond their first week and 44% remained beyond their first month. Even after 200 days, 15% of participants remained consistent contributors. The exceptionally high rate of retention likely reflects the easy-to-use app design, as well as the dedication of participants to contribute to answering a question of great personal interest to them.

In all, there were 445 days available for study. The top 10% (45 days) with the highest percentage of participants having a +1 or greater pain event were termed HIGH. The bottom 10% with the lowest percentage of participant experienced within the United Kingdom.

In a previous study, we performed an independent analysis of the dataset using a case-crossover design, an epidemiological method that compares, for each individual, days with a pain event to control days without a pain event, thus controlling for all time-invariant participant characteristics. The analysis, based on a much smaller dataset, demonstrated statistically significant—albeit modest—relationships between high pain events and high relative humidity, low sea level pressure, and high wind speed. Temperature and precipitation did not have statistically significant relationships with pain. These results held up even when accounting for mood and physical activity.

Here, we see if these results are reproducible from a meteorological perspective, using synoptic climatology and compositing to visualize the synoptic-scale weather patterns associated with pain.



Synoptic composites of HIGH and LOW days. Anomalies computed relative to the weighted average of the daily 1981–2010 means.

participants having a +1 or greater pain event were termed LOW. The most painful days (HIGH) had 23% of participants across the United Kingdom reporting an increase in pain, and the least painful days (LOW) had only 10% of participants reporting an increase in pain. Hourly weather observations were averaged to produce daily weather conditions for each individual who submitted daily pain reports as well as U.K.-average weather for the day.

Relating pain to weather

Most HIGH days occurred from January to June 2016, whereas most LOW days occurred from June 2016 to January 2017. Variations by month in the fraction of participants experiencing a pain event suggest that seasonal changes in weather might affect levels of pain, although we could not discount that the population of active participants changed over the course of the study.

The 1200 UTC weather maps for each HIGH and LOW day were averaged together to produce synoptic composites. Detailed composite analysis of 500-hPa geopotential height, sea level pressure, precipitation, 925-hPa moisture, wind, and temperature reveal remarkably different weather patterns during HIGH versus LOW days.

On HIGH days, the jet stream was aimed right at the United Kingdom. This is associated with below-normal (or low) pressure, and thus more wind, moisture, and precipitation. In contrast, on LOW days, the jet stream tended to blow north of the United Kingdom, bringing above-normal (or high) pressure and associated below-normal humidity and precipitation rates, and weaker winds. In short, the results provide the strongest scientific support for people who have been saying that the weather affects their pain but who have not been taken seriously by other people, their own doctors included.

Discussion

Although the results are intriguing, they apply only to the United Kingdom. Similar studies should be replicated elsewhere to see if these findings can be generalized. Another caveat is that many participants had multiple health conditions causing pain, and even those with the same underlying condition may have felt pain differently. We didn't group the participants into different disease categories in our analysis—a topic of future research. Also, untangling the effect of those who stayed inside—due to severe pain or other reasons—requires further analysis.

Although specific weather patterns may not be the primary cause of people's pain, our results demonstrate that weather does modulate pain in at least some individuals. Who is most susceptible remains to be determined.

People have been talking about the effect of weather on their pain for millennia, so why is this particular research project important? One reason is that our study had both the longest duration and the most participants of all previous studies, allowing greater confidence in the fidelity of our results. Another reason is that this research compares for the first time the weather patterns on days with a large number of people reporting pain and days with a low number reporting pain. Finally, our research begins to shed light on the environmental conditions that modulate pain, insight that might be explored further for improving the treatment, management, and forecasting of pain.

METADATA

BAMS: What would you like readers to learn from this article?

David M. Schultz (University of Manchester): *Three-quarters of people who suffer from pain report that they believe that weather influences their pain—yet their concerns are often dismissed by others, including their doctors. Although weather may not be the primary influence on people's pain, we show that it can modulate pain events in some people. Our study gives patients more support for their beliefs, showing that their beliefs have validity.*

BAMS: How did you get involved in this project?

DMS: *Project lead Prof. William Dixon had heard anecdotal reports of the influence of weather on pain from his patients in clinic. He was working on using mobile technology to better collect information on patient's symptoms in between visits to the clinic when he realized that the GPS sensor in mobile phones could be linked to the closest weather station to record weather data closest to the patients. A collaboration with uMotif, a company that creates mobile health*

Some of the main research team on the Cloudy With a Chance of Pain project: Ana M. Vicedo-Cabrera, Antonio Gasparrini, Mark Lunt, Anna L. Beukenhorst, William G. Dixon, David M. Schultz, Malcolm Maclure (on screen), John McBeth, Belay B. Yimer, and Ricardo Silva.

apps, led to the app to collect patient data. Thus, William could envision a research project that would test out ideas about engagement with health apps and answer a longstanding question about the effect of weather on pain. A chance meeting between William and my Head of Department brought me on board.

BAMS: What attracted you to the idea of relating weather to chronic pain?

DMS: *I had a latent interest in weather and pain because both of my parents have arthritis. I jumped at the opportunity to get involved in the project. It's been one of the most satisfying research projects of my career.*

BAMS: What surprised you the most about the work you document in this article?

DMS: *The enthusiasm of the citizen scientists who entered their data on a daily basis. When we designed the project, we had hoped we might get 1,000 people to sign up. However, William appeared on the BBC's* Trust Me I'm a Doctor *and on* BBC Breakfast. *Because of the media interest, and the enthusiastic and financial support of Versus Arthritis (formerly Arthritis Research UK), we had over 13,000 people download the app. Shortly after William's appearance on* BBC Breakfast, *people were signing up to the study at a rate of one per second, testing the capacity/resilience of the server at uMotif. And these people were enthusiastic about participating! Most people download apps, use them a few times, and then stop. Other studies of engagement with mobile apps confirm such low retention rates of participants with apps. However, even after six months into our study, 15% of participants were still entering their data nearly every day. The success of this study is due to their dedication. We are so grateful for their participation.*

BAMS: What about surprises in the results?

DMS: *We had expected a weak relationship and maybe some insight into the weather–pain relationship. The strength of the relationship and its robustness were surprising.*

BAMS: What was the biggest challenge in the project?

DMS: *The principal challenge was rapidly scaling up our project from a small funded pilot study of 20 people to a national-scale project. Although we were excited that the opportunity to be on BBC television would help promote our project nationally and boost recruitment, we were challenged by finding a source of funding to build the app and pay for data collection and hosting on such a short time scale. That meant from the time of recording the show to when it aired, we had six months to get funding for a national-scale project. Fortunately, Versus Arthritis was open to considering a proposal with such a short time frame, and they provided enough funding from them to carry out the data collection. The issue of data analysis was another story. In that case, other synergies (e.g., incoming Ph.D. student and postdoctoral researcher, collaborations with other researchers at other institutions) happened later that allowed us to bring a diverse group of people onto the team who could help analyze the data.*

General Section
Research Paper

OPEN

PAIN REPORTS

Heterogeneity in the association between weather and pain severity among patients with chronic pain: a Bayesian multilevel regression analysis

Belay B. Yimer[a,b,*], David M. Schultz[c,d], Anna L. Beukenhorst[a,e], Mark Lunt[a,b], Huai L. Pisaniello[a,f], Thomas House[g], Jamie C. Sergeant[a,h], John McBeth[a,b], William G. Dixon[a,b]

Abstract

Introduction: Previous studies on the association between weather and pain severity among patients with chronic pain have produced mixed results. In part, this inconsistency may be due to differences in individual pain responses to the weather.

Methods: To test the hypothesis that there might be subgroups of participants with different pain responses to different weather conditions, we examined data from a longitudinal smartphone-based study, Cloudy with a Chance of Pain, conducted between January 2016 and April 2017. The study recruited more than 13,000 participants and recorded daily pain severity on a 5-point scale (range: no pain to very severe pain) along with hourly local weather data for up to 15 months. We used a Bayesian multilevel model to examine the weather–pain association.

Results: We found 1 in 10 patients with chronic pain were sensitive to the temperature, 1 in 25 to relative humidity, 1 in 50 to pressure, and 3 in 100 to wind speed, after adjusting for age, sex, belief in the weather–pain association, mood, and activity level. The direction of the weather–pain association differed between people. Although participants seem to be differentially sensitive to weather conditions, there is no definite indication that participants' underlying pain conditions play a role in weather sensitivity.

Conclusion: This study demonstrated that weather sensitivity among patients with chronic pain is more apparent in some subgroups of participants. In addition, among those sensitive to the weather, the direction of the weather–pain association can differ.

Keywords: Chronic pain, Weather, Musculoskeletal diseases, Multilevel modelling, Observational studies

1. Introduction

There is a strong belief among patients with chronic pain that pain severity is influenced by the weather.[17,23] However, studies investigating the association between weather and pain have yielded conflicting results.[2,21] One possibility for this lack of consensus is that some people within the population are highly sensitive to the weather, others are less sensitive, and some are not sensitive.[16,21,23] Such differences among individuals, including the subjective and highly personal nature of pain experience, are well known.[11] These individual differences are not merely a byproduct of idiosyncrasies in the reporting of pain but may be a result of interindividual differences in cerebral activation evoked by the same painful stimulus.[6,9] However, most of the previous studies have focused on the average effect of weather on pain severity at a population level and have not investigated individual differences.[8,25] Understanding individual variation and

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] *Centre for Epidemiology Versus Arthritis, University of Manchester, Manchester, United Kingdom,* [b] *NIHR Greater Manchester Biomedical Research Centre, Manchester Academic Health Science Centre, University of Manchester, Manchester, United Kingdom,* [c] *Department of Earth and Environmental Sciences, Centre for Atmospheric Science, University of Manchester, Manchester, United Kingdom,* [d] *Centre for Crisis Studies and Mitigation, University of Manchester, Manchester, United Kingdom,* [e] *Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA,* [f] *Discipline of Medicine, The University of Adelaide, Adelaide, Australia,* [g] *School of Mathematics, The University of Manchester, Manchester, United Kingdom,* [h] *Centre for Biostatistics, University of Manchester, Manchester, United Kingdom*

**Corresponding author. Address: Centre for Epidemiology Versus Arthritis, University of Manchester, Stopford Building, Manchester M13 9PL, United Kingdom. Tel.: +44 (0)161 306 6000. E-mail address: belaybirlie.yimer@manchester.ac.uk (B.B. Yimer).*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painrpts.com).

PR9 7 (2022) e963

http://dx.doi.org/10.1097/PR9.0000000000000963

the factors contributing to individual differences in pain may provide insights into pain mechanisms. However, assessing individual variation requires fitting models explicitly designed to account for individual-specific responses and their associated uncertainty intervals. Without repeated observations of the same individuals over a sufficient time, this is not possible. Two studies attempted to model weather–pain association at an individual level using a multilevel modelling framework that explicitly models individual-level heterogeneity.[3,10] However, the small sample size and limited follow-up hampered the robustness of their analysis. As a result, there is no robust evidence for heterogeneity in the weather–pain association.

Recently, we conducted a large UK-based smartphone study, Cloudy with a Chance of Pain (www.cloudywithachanceofpain.com), recruiting more than 13,000 patients across the UK over 15 months.[7] Participants with a range of underlying pain conditions tracked their daily symptoms through the study smartphone application (app) for 6 months or more while the GPS in the smartphone enabled local weather data collection. An analysis of the Cloudy with a Chance of Pain data set by Dixon et al.[7] demonstrated higher relative humidity and wind speed and lower atmospheric pressure were associated with increased pain severity. The analysis used a case-crossover method to generate population-level estimation of the weather–pain association that corrected for the individual difference in unmeasured baseline factors. However, the Cloudy with a Chance of Pain data set also provided a unique opportunity for exploring individual-level heterogeneity. In this study, we test the hypothesis that there is an association between the weather and pain severity that is only apparent in a subgroup of participants. We then examine the extent to which the difference in underlying pain condition captures individual heterogeneity.

2. Methods

2.1. Design and study sample

Cloudy with a Chance of Pain (www.cloudywithachanceofpain.com)[7] was conducted between January 20, 2016, and April 20, 2017, to understand the relationship between weather and pain. A total of 13,207 users across the UK over the 12-month recruitment period downloaded the study smartphone application. Information including age, sex, underlying pain condition, and participants' belief about the weather–pain association ("How likely do you think it is that the weather is associated with pain?" measured on a 1- to 10-point scale with 1 being "not at all likely" and 10 being "extremely likely") was recorded at baseline. Participants were requested to submit their pain severity level on an ordinal scale with 5 categories and 9 other variables, including mood, activity, and fatigue, daily. Participants were followed up from their first pain severity level entry up to the last pain severity level entry. A total of 10,584 participants had completed baseline information and at least one pain entry, with 6850 (65%) participants remaining in the study beyond their first week.[7] A detailed description of the study is presented in the research conducted by Dixon et al.[7] Ethical approval was obtained from the University of Manchester Research Ethics Committee (ref: ethics/15522) and from the NHS IRAS (ref: 23/NW/0716).

2.2. Inclusion criteria

Participants were included in the final cohort for this analysis if they fulfilled the following criteria: (1) downloaded the app, (2) provided consent, (3) completed the baseline questionnaire, and (4) contributed at least 2 days of pain severity data.

2.3. Primary outcome measure

The outcome of interest was the daily self-reported pain severity level recorded on a 1- to 5-point ordinal scale (1: no pain, 2: mild pain, 3: moderate pain, 4: severe pain, and 5: very severe pain). Participants were asked to report self-reported pain severity level every day using the smartphone application, prompted by a daily notification at 6:24 PM.

2.4. Exposures

The exposures of interest in this study were 4 state weather parameters, namely the average daily temperature, pressure, relative humidity, and wind speed a participant was exposed to each day. Study participants' locations were recorded at each hour of the day using the study app. Weather information, including temperature, pressure, relative humidity, and wind speed, was retrieved by linking participants' locations to the nearest Met Office weather station. When participants were outside the UK during the study period, their data were not analysed because we were unable to link to non-UK weather stations.

2.5. Covariates

We considered age (in years), sex, baseline beliefs about the association between weather and pain, the daily record of mood (on a 1- to 5-point scale; 1: depressed, 2: feeling low, 3: not very happy, 4: quite happy, and 5: very happy) and exercise (on a 1- to 5-point scale; 1: no exercise, 2: less than 30 minutes of light activity, 3: 30+ minutes of light activity, 4: less than 30 minutes of strenuous activity, and 5: 30+ minutes of strenuous activity) as possible factors[7] that may influence the weather–pain association.

2.6. Statistical analysis

The association between weather and pain severity was tested with a multilevel ordinal probit model.[14,18,24] This model was considered ideal for this analysis because it allowed an estimate of the average response across the group through the fixed-effect terms, and it could explicitly model participant-level heterogeneity using random-effect terms. The model allows every participant in our study to have their unique response to the weather while also improving population-average estimates by pooling information across participants.[13] The model also appropriately handles the primary outcome's ordinal scale and irregular (ie, unbalanced) repeated measurements.[20]

We developed a multivariable multilevel model that included the 4 state weather parameters adjusted for age, sex, belief, time since entry to the study, mood, and exercise as fixed effects. A linear relationship was assumed for all variables in the model except for time since entry to the study, which was modelled nonparametrically using a cubic spline[26] (ie, the data entirely determined the relationship between time and the response). We assumed a linear relationship between weather parameters and pain severity because the complex nonparametric relationship did not produce a better fit. We included time since entry to the study in the model as a means of filtering out unmeasured time-varying factors that may influence a participant's pain severity reports. In addition to the abovementioned fixed-effect terms, the model included 5 correlated participant-specific random effects, namely a random intercept and a random effect for each of the 4 state weather parameters. The random intercept term captures

the between-participant variation not explained by the baseline factors. The random effects for the weather parameters allow the weather effects to vary over study participants.

The Bayesian estimation approach assuming noninformative priors,[15] described in detail in Section 2 of the supplementary file (available at http://links.lww.com/PR9/A145), was followed to estimate model parameters. We examined the trace plots and the posterior distribution plot and performed posterior predictive checks[12] to assess model convergence. We reported the estimated regression coefficients (β) along with their associated 95% credible intervals. The regression coefficients represent the change in the *z* score or probit index for a one-unit change in the weather parameter. To quantify the change in the predicted probability for each pain severity response level for a one-unit change in the weather parameter, we used a summary measure called marginal effect at the mean.[1] The marginal effect at the mean represents the marginal effect of the explanatory variables of interest while holding the other variables in the model at their respective mean values. Furthermore, we presented the participant-level regression coefficients with a 95% credible interval as a forest plot along with the population-level effects (β). We divided participants into groups that are statistically distinct from one another by determining whether their credible intervals overlap.[19] The prevalence of the participant's underlying diagnosis was then compared between groups, with the data presented as a bar plot. Owing to statistical power issue, we were unable to perform a statistical test on the difference in weather sensitivity by the participant's underlying diagnosis. We used the R package brms[4] based on Stan[5] to fit the model. The R source code is made available at belayb/Cloudy-Probit: Bayesian Multilevel analysis (github.com).

3. Results

3.1. Participants' characteristics

Of the 13, 000 participants recruited for the study, a total of 6213 participants who had completed baseline information, submitted at least 2 days of pain reports, and had hourly location data sufficient to retrieve complete weather information to produce daily means were included in the analysis. Study participants included in the analysis had a mean age of 49 years (SD: 13.0); most of them were female individuals (82%), and most of the participants believed in an association between weather and their pain (median score 7 of 10, interquartile range [IQR]: 6–9) (**Table 1**). Approximately 35% of the participants experienced unspecified arthritis, followed by osteoarthritis (29%) and fibromyalgia (27%) (**Table 1**). The characteristics of those included in the analysis were similar to the full cohort of participants (Table S1 in the supplementary file, available at http://links.lww.com/PR9/A145). The participants included in the analysis were followed up for a median of 106 days (IQR: 53–215). On average, they contributed pain severity data for 65% of the days during the period when they were actively contributing data to the study. Overall, participants tended to report mild or moderate pain (2 or 3 on our 5-point scale) approximately 70% of the time (Figure S1 in the supplementary file, available at http://links.lww.com/PR9/A145).

Table 1

Baseline characteristics of study participants.

Characteristics	Final cohort (N = 6213)
Demographics	
Female, N (%)	5519 (82.4)
Age, mean (SD)	48.68 (13.0)
Diagnosis, N (%)*	
Arthritis (type not specified)	2135 (34.4)
Osteoarthritis	1797 (28.9)
Fibromyalgia/chronic widespread pain	1707 (27.5)
Rheumatoid arthritis	1176 (18.9)
Neuropathic pain	975 (15.7)
Chronic headache (including migraine)	630 (10.1)
Ankylosing spondylitis/spondyloarthropathy	552 (8.9)
Gout	213 (3.4)
Other/no medical diagnosis	1179 (19.0)
Belief in weather–pain association	
Belief that the weather influences pain on a scale of 1–10, median (IQR)	7 (6–9)

* Participants may report more than one pain condition, and when they do, they are counted multiple times in the abovementioned table.
IQR, interquartile range.

3.2. Population-level weather–pain association

Table 2 summarizes parameter estimates and their associated 95% credible intervals for the Bayesian multilevel ordinal probit model. At a population level, participants exposed to high relative humidity (0.041, 95% CI: 0.034–0.048) or high wind speed (0.012, 95% CI: 0.009–0.014) have a higher likelihood of experiencing a higher level of pain (**Table 2**). Similarly, participants exposed to low temperatures (−0.003, 95% CI: −0.005 to −0.001) or low pressures (−0.010, 95% CI: −0.015 to −0.005) have a higher likelihood of experiencing a higher level of pain. That is, an increase in relative humidity by 10 percentage points increases the probability of reporting moderate pain or above by 1.5%, and an increase in wind speed by 1 $m{\cdot}s^{-1}$ increases the probability of reporting moderate pain or above by 0.40%. Similarly, an increase in temperature by 1°C decreases the probability of reporting moderate pain or above by 0.1%. An increase in pressure by 10 mbar decreases the probability of reporting moderate pain or above by 0.4% (**Table 2**). In general, the population level estimated that weather–pain association for all considered weather parameters were modest.

3.3. Exposure effect heterogeneity

We evaluated the participant-level weather–pain associations to identify subgroups within the population who were sensitive to the weather. **Figure 1** shows the estimated participant-level regression coefficients and their associated credible intervals for each weather parameter ranked by their median values. We divided participants into groups that are statistically distinct from one another by determining whether their credible intervals overlap. For each of the considered weather parameters, this method identifies 2 distinct clusters (both coloured blue) and a third cluster (coloured grey) for the participants who cannot be statistically distinguished from the members of the 2 distinct clusters. These 3 clusters are given names based on the direction of the weather–pain association: low-value sensitive (ie, participants with negative posterior credible intervals), high-value sensitive (ie, participants with positive posterior credible intervals), and undetermined (ie, participants with credible intervals that overlap with zero). In this study, we considered the name not sensitive instead but settled on undetermined because lack of statistical significance does not mean the absence of a relationship.

Most of the participants belonged to the undetermined cluster for all weather parameters, implying that, for most of the participants, there is not enough evidence to indicate that they possess sensitivity to the weather–pain association. The size of the low-value sensitive and high-value sensitive clusters varies by weather parameter (**Figure 1**). For example, there

Table 2

Association between weather and pain—parameter estimates from the Bayesian multilevel ordinal probit model.

Weather parameters	Estimate (β)*	95% credible interval	Marginal effects at mean (MEM)†
Temperature (per 1°C)	−0.003	(–0.005 to −0.001)	−0.001
Pressure (per 10 mbar)	−0.010	(–0.015 to −0.005)	−0.004
Relative humidity (per 10%)	0.041	(0.034 to 0.048)	0.015
Wind speed (per 1 $m \cdot s^{-1}$)	0.012	(0.009 to 0.014)	0.004

* The model is adjusted for age (in years), sex, belief, mood, and exercise.
† MEM represents the change in probability of experiencing moderate pain or above as the weather parameter value increases.

were a similar proportion of participants for whom relatively lower temperature was associated with a higher level of pain (6.3%) as there were participants (4.7%) for whom the higher temperature was associated with an increase in their pain, resulting in a very modest overall effect of temperature. On the other hand, the participant-level regression coefficients of relative humidity and wind speed were skewed to the right of zero. More participants (2.9% for relative humidity and 2.2% for wind speed) were sensitive to higher values of these weather parameters than to lower values (0.6% for relative humidity and 0.6% for wind speed). Similarly, proportionally more participants (1.6%) were sensitive to low pressure than high pressure (0.7%). Most of the participants (72.5%) classified as weather sensitive possessed sensitivity to a single weather parameter (Figure S3 in the supplementary material, available at http://links.lww.com/PR9/A145).

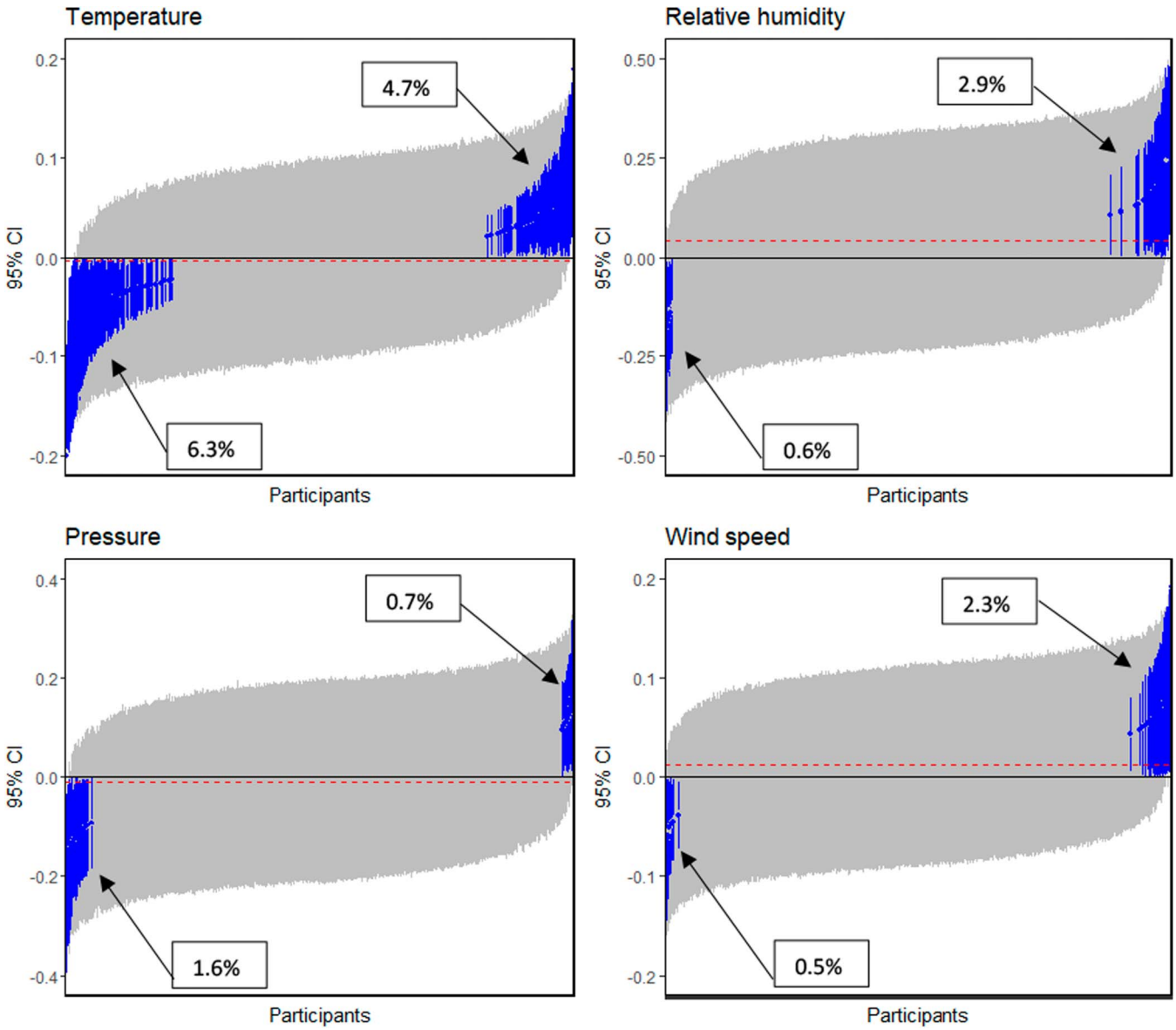


Figure 1. Heterogeneity in weather–pain association. The 95% credible interval for each of the 4 estimated weather effects for each participant sorted by their median values of the estimated effect sizes. Effect sizes are on the latent scale. Intervals shown in blue do not cross zero. The horizontal dotted red line is the population average weather effect on pain severity, consistent with the population-level result listed in Table 2.

To understand the role of the participant's underlying disease diagnosis on weather sensitivity, we explored the distribution of the weather sensitivity group in each of the pain conditions of participants. **Figure 2** presents the distribution of low-value sensitive and high-value sensitive clusters by participant's disease diagnosis for each of the 4 weather parameters. To simplify the analysis, we grouped various diseases into one of the 4 categories: osteoarthritis, fibromyalgia or chronic widespread pain, inflammatory arthritic pain (rheumatoid arthritis and ankylosing spondylitis or spondyloarthropathy), and other chronic pain. For clarity, we considered only participants with a single diagnosis (n = 3355) in **Figure 2**. Based on visual inspection of **Figure 2**, there are for the most part no major differences in the prevalence of weather sensitivity observed between participant's disease diagnosis, although there is perhaps a hint that participants with inflammatory arthritis are more commonly sensitive to low temperatures and have a greater differential sensitivity to pressure.

4. Discussion

4.1. Summary of principal findings

This study tested the hypothesis that there is an association between the weather and pain severity that is only apparent in a subgroup of participants using a large longitudinal data set. The data presented in this study support that hypothesis. After adjusting for age, sex, belief in the weather–pain association, mood, and activity level for each of the 4 weather parameters considered (ie, the average daily temperature, pressure, relative humidity, and wind speed a participant was exposed to each day), we identified 3 statistically distinct clusters of patients with chronic pain who were each influenced by the weather differently: low-value sensitive, high-value sensitive, and undetermined. Eleven percent of participants were sensitive to temperature, of which 6.3% were sensitive to low temperature. On the other hand, most of those sensitive to relative humidity and wind speed were high-value sensitive (2.9% of 3.5% and 2.3% of 2.8%, respectively). Similarly, most of those sensitive to pressure were low-value sensitive (1.6% of 2.3%).

This study also examined the role of the underlying conditions (ie, participant's disease diagnosis) on their sensitivity to the weather. There is no definite indication of individual underlying pain conditions explaining individual-specific weather–pain association, although participants with inflammatory arthritis may have been more sensitive to cold than the other conditions.

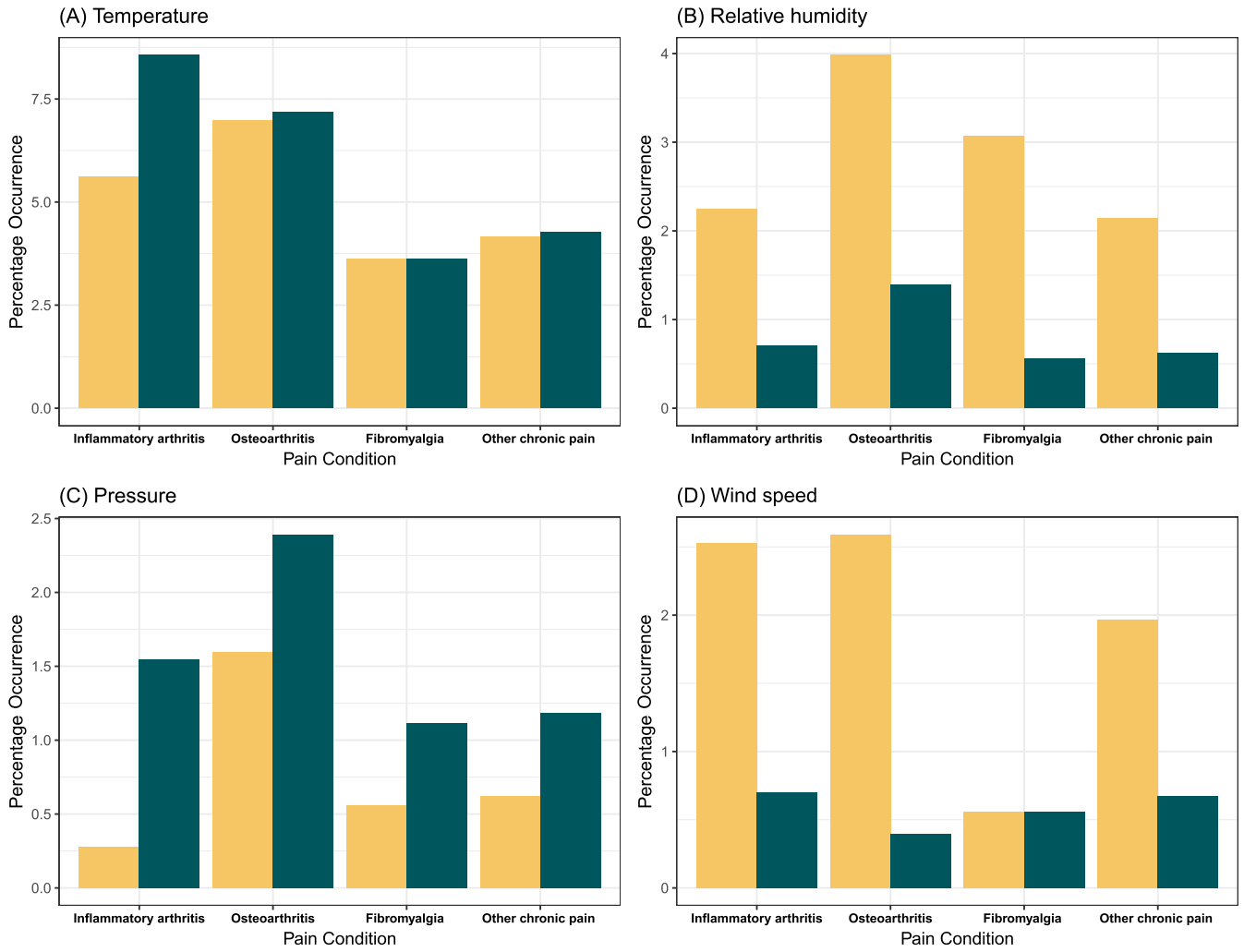


Figure 2. Distribution of weather sensitivity group by underlying pain conditions of participants. Rheumatoid arthritis and ankylosing spondylitis or spondyloarthropathy are grouped as inflammatory arthritic pain. Yellow bars represent high-value sensitive clusters, and green bars represent low-value sensitive clusters.

4.2. Methodological strengths

This study sets individual variation at the forefront and aims to quantify the influence of weather on pain severity at the individual level. Previous studies focus on an average effect of weather on pain severity at a population level.[8,25] In the presence of individual heterogeneity, the weak weather–pain association at a population level based on this approach does not rule out the possibility of a stronger weather–pain association at an individual level. Indeed, the lack of a population-level association might be because the study participants are composed of about an equal number of participants affected by the weather in opposite directions, thereby cancelling out the effect in the population as a whole. In the case of a statistically significant positive or negative association at the population level, and in the presence of heterogeneity in the weather–pain relationship, it is problematic to use the resulting population estimates to provide clinical advice to an individual patient because the population-level estimates may not meaningfully apply to individuals. This issue underscores the importance of explicitly modelling individual-level heterogeneity. However, without repeated observations of the same individuals over a long period, it is impossible to quantify individual-level heterogeneity and identify subgroups that behave differently.

This study uses a large data set obtained from the Cloudy with a Chance of Pain study,[7] which produced a unique data set by recruiting more than 13,000 participants with sustained daily self-reported data and accurate weather information to which they were exposed over many months. The data set also recorded daily self-reported mood and activity level, which were ideal for estimating the weather effect that was acting not through these variables. We used a multilevel modelling approach to analyse the data and investigate the influence of weather on pain severity. The modelling approach allows every participant in the study to have their unique response to the weather while also improving the population-level average estimate by pooling information across study participants.[20] When estimating the average effect of weather on the population, the multilevel modelling approach prevented oversampled individuals from unfairly dominating the result by considering the differential uncertainty across participants.[20]

A limitation in our study was that our study participants were aware of the study objective, which may raise possible information bias where observed weather could influence participants' symptom reporting. However, our analysis has been adjusted for previous belief, and hence, information bias will not fully explain the observed association. Also, the findings from this study cannot necessarily be extrapolated to different climates where the weather is different.

4.3. Comparison with other studies

Previous studies have reported a relatively higher percentage of weather-sensitive individuals. For example, Fagerlund et al.[10] investigated individual differences in weather sensitivity using a multilevel modelling framework. They found significant individual differences, with a subgroup of patients (20%) behaving contrarily to most patients by reporting increased pain with increased atmospheric pressure. Similarly, Bossema et al.[3] used a multilevel modelling approach to investigate individual heterogeneity. They reported a positive association between the weather variables (ie, temperature, sunshine duration, perception, pressure, and relative humidity) and pain in approximately one-third of the patients, a negative association in one-third of the patients, and no association in the remaining patients. The 2 studies considered only patients with fibromyalgia. Compared with our result, the higher reported percentage may be attributed to the difference in the methodology. For example, Bossema et al.[3] used Pearson correlation between the fibromyalgia symptom and the weather condition for each patient to identify individual-level association rather than using the multilevel model used to estimate population effect in their analysis.

Two studies[16,22] among others[2] examined subgroups sensitive to the weather by analysing each participant's data individually and reported substantial difference among individuals in weather sensitivity. However, such analysis is prone to overfitting and may lead to spurious associations.

4.4. Implications of the study

This study demonstrated that weather sensitivity among patients with chronic pain is a phenomenon more apparent in some subgroups of participants. In addition, among those sensitive to the weather, the direction of the weather–pain association can differ. When considering future potential benefits and applications of understanding the association between weather and pain, such as developing a "pain forecast" to help patients predict their forthcoming pain, our results would support the need for a personalised prediction.

Disclosures

W.G. Dixon has received consultancy fees from Google and AbbVie unrelated to this study.

Acknowledgements

Cloudy with a Chance of Pain was supported by Versus Arthritis (grant reference 21225). This research was further supported by the Centre for Epidemiology Versus Arthritis (grant number 21755) and the NIHR Manchester Biomedical Research Centre.
A.L. Beukenhorst is supported by the Medical Research Council (grant number MR/N013751/1). Data are available on request.

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PR9/A145.

Article history:
Received 10 June 2021
Received in revised form 25 August 2021
Accepted 26 August 2021

References

[1] Agresti A, Tarantola C. Simple ways to interpret effects in modeling ordinal categorical data. Stat Neerlandica 2018;72:210–23.
[2] 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;161:668–83.
[3] Bossema ER, van Middendorp H, Jacobs JW, Bijlsma JW, Geenen R. Influence of weather on daily symptoms of pain and fatigue in female patients with fibromyalgia: a multilevel regression analysis. Arthritis Care Res (Hoboken) 2013;65:1019–25.
[4] Bürkner P-C. brms: an R package for Bayesian multilevel models using Stan. J Stat Softw 2017;80:1–28.
[5] Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A. Stan: a probabilistic programming language. J Stat Softw 2017;76:1–32.

[6] Coghill RC, McHaffie JG, Yen YF. Neural correlates of interindividual differences in the subjective experience of pain. Proc Natl Acad Sci U S A 2003;100:8538–42.
[7] Dixon WG, Beukenhorst AL, Yimer BB, Cook L, Gasparrini A, El-Hay T, Hellman B, James B, Vicedo-Cabrera AM, Maclure M, Silva R, Ainsworth J, Pisaniello HL, House T, Lunt M, Gamble C, Sanders C, Schultz DM, Sergeant JC, McBeth J. How the weather affects the pain of citizen scientists using a smartphone app. NPJ Digit Med 2019;2:105–9.
[8] Duong V, Maher CG, Steffens D, Li Q, Hancock MJ. Does weather affect daily pain intensity levels in patients with acute low back pain? A prospective cohort study. Rheumatol Int 2016;36:679–84.
[9] Emerson NM, Zeidan F, Lobanov OV, Hadsel MS, Martucci KT, Quevedo AS, Starr CJ, Nahman-Averbuch H, Weissman-Fogel I, Granovsky Y, Yarnitsky D, Coghill RC. Pain sensitivity is inversely related to regional grey matter density in the brain. PAIN 2014;155:566–73.
[10] Fagerlund AJ, Iversen M, Ekeland A, Moen CM, Aslaksen PM. Blame it on the weather? The association between pain in fibromyalgia, relative humidity, temperature and barometric pressure. PLoS One 2019;14: e0216902.
[11] Fillingim RB. Individual differences in pain: understanding the mosaic that makes pain personal. PAIN 2017;158(suppl 1):S11.
[12] Gabry J, Simpson D, Vehtari A, Betancourt M, Gelman A. Visualization in Bayesian workflow. J R Stat Soc Ser A (Statistics Society) 2019;182: 389–402.
[13] Gelman A. Multilevel (hierarchical) modeling: what it can and cannot do. Technometrics 2006;48:432–5.
[14] Gelman A, Hill J. Data analysis using regression and multilevel/ hierarchical models. Cambridge, United Kingdom: Cambridge University Press, 2006.
[15] Gelman A, Jakulin A, Pittau MG, Su Y-S. A weakly informative default prior distribution for logistic and other regression models. Ann Appl Stat 2008; 2:1360–83.
[16] Guedj D, Weinberger A. Effect of weather conditions on rheumatic patients. Ann Rheum Dis 1990;49:158–9.
[17] Hagglund KJ, Deuser WE, Buckelew SP, Hewett J, Kay DR. Weather, beliefs about weather, and disease severity among patients with fibromyalgia. Arthritis Care Res 1994;7:130–5.
[18] Lee Y, Nelder JA. Hierarchical generalized linear models. J R Stat Soc Ser B (Methodological) 1996;58:619–56.
[19] Lum K, Dunson DB, Johndrow J. Closer than they appear: a Bayesian perspective on individual-level heterogeneity in risk assessment. 2021. arXiv preprint arXiv:2102.01135.
[20] McElreath R. Statistical rethinking: a Bayesian course with examples in R and Stan. CRC Press, 2020.
[21] Smedslund G, Hagen KB. Does rain really cause pain? A systematic review of the associations between weather factors and severity of pain in people with rheumatoid arthritis. Eur J Pain 2011;15:5–10.
[22] Smedslund G, Mowinckel P, Heiberg T, Kvien TK, Hagen KB. Does the weather really matter? A cohort study of influences of weather and solar conditions on daily variations of joint pain in patients with rheumatoid arthritis. Arthritis Rheum 2009;61:1243–7.
[23] Timmermans EJ, Van Der Pas S, Schaap LA, Sánchez-Martínez M, Zambon S, Peter R, Pedersen NL, Dennison EM, Denkinger M, Castell MV, Siviero P, Herbolsheimer F, Edwards MH, Otero A, Deeg DJ. Self-perceived weather sensitivity and joint pain in older people with osteoarthritis in six European countries: results from the European Project on OSteoArthritis (EPOSA). BMC Musculoskelet Disord 2014;15: 66.
[24] Varin C, Czado C. A mixed autoregressive probit model for ordinal longitudinal data. Biostatistics 2009;11:127–38.
[25] Wilder FV, Hall BJ, Barrett JP. Osteoarthritis pain and weather. Rheumatology (Oxford) 2003;42:955–8.
[26] Wood SN. Generalized additive models: an introduction with R. CRC press, 2017.

Appendix E: Lay Summary of Overseas Fellowship – Mobile Health Applications for Monitoring Musculoskeletal Conditions (Arthritis Australia)

Explore 2019 Research

Oral health and rheumatoid arthritis

Ky-anh Nguyen

3D printing of bio-glue to repair cartilage

Dr Claudia Di Bella

Falls prevention for osteoarthritis care

Prof Ilana Ackerman

Balancing muscle force and persistent knee pain in adolescents

Dr Kylie Tucker

Knee cap pain in young and middle aged adults

Dr Natalie Collins

The burden of cancer in systemic sclerosis

Dr Kathleen Morrisroe

Mobile Health applications for monitoring musculoskeletal conditions

Dr Huai Leng (Jessica) Pisaniello (nee Yong)

Recipient:	Dr Huai Leng (Jessica) Pisaniello (nee Yong)
Intended department:	The Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, School of Biological Sciences, Faculty of Biology, Medicine Health- The University of Manchester
Project:	*The Role of Mobile Health application in real time capture of self-reported symptoms and longitudinal activity, and its feasibility in patient-focused remote monitoring in musculoskeletal*

	disorders.

Mobile health (mHealth) implementation in clinical care and research has great potential to capture real-time data longitudinally, when compared to traditional epidemiological methods. At present, self-reported symptoms or experience can be captured conveniently and stored digitally within mHealth technologies (e.g. smartphone, smartwatch) and mobile applications (apps). These data, although structurally heterogeneous, are rich sources of health-related information that could reflect a patient's health journey.

In people with arthritis and other musculoskeletal diseases, tracking symptoms using mobile health apps can greatly improve our understanding of how these diseases affect people and how symptoms can change from day to day. The burden of chronic pain in arthritis is huge. We know that pain remains one of the most important, and yet, a challenging symptom to manage and treat in patients with musculoskeletal diseases. Chronic pain is a dynamic process and can be unpredictable, particularly for those with underlying inflammatory musculoskeletal diseases and concomitant chronic pain condition. Patients are often asked to summarise their overall pain severity since the last clinic visit, which could be weeks to months. This averaged pain severity may not necessarily reflect the overall pain severity over time and in particular, the fluctuation of pain over time. Under- or over-estimation of pain experience can have a huge impact on treatment decision, especially with biologic prescription and analgesic choice. Longitudinal capture of patient's pain symptoms will help the clinicians in clarifying the definition of disease 'flare' and in guiding the trajectory of pain management. From the patients' perspectives, being able to have better understanding of their pain patterns and levels of pain variability will allow greater sense of control in managing their pain. My research aims to understand how mobile health is implemented in collecting patient- generated health data, as well as how these temporally rich data are analysed. Specifically, my research aims to examine the long-term and short-term day-to-day pain variability over time in a chronic pain study cohort.

As the recipient of this prestigious overseas fellowship grant year 2018, I had the opportunity to work at the Centre for Epidemiology Versus Arthritis in Manchester, UK under the supervision of Professor William Dixon. This 15-month fellowship in Manchester has set off to an exceptionally invaluable foundation for my first year of PhD. I was working with Professor Dixon and the *Cloudy* team on a large, UK nationwide, prospective mobile health study called *Cloudy with a Chance of Pain*. This study, which was conducted in January 2016, aimed to examine the association between weather and pain in people living with chronic pain. Although I was not involved in the setup of this project, I came to learn from many in the *Cloudy* team about the success of this study, in particular the patient and public involvement and recruitment strategies.

The process of preparing and analysing such a large data set has been a huge undertaking, only made possible with expertise input from many researchers with different background (rheumatologists, epidemiologists (national and international), meteorologist, statisticians, mathematician, PhD student with health informatics background and project manager). As a newly trained rheumatologist, doing this fellowship has certainly honed my skills and knowledge in applied epidemiology and statistics in musculoskeletal research.

In Manchester, as a clinical research fellow in the department, I managed to attend various in- house departmental courses, lectures and seminars. These include a 6-month course on epidemiology, genetic epidemiology and statistics in the first half of year, a 3-month course on 'Statistical Modelling with Stata' by Dr Mark Lunt, weekly departmental seminars, monthly applied epidemiology sessions, CfE scientific meetings and journal club. I was able to attend the first Digital Epidemiology Summer School course in July 2018, led by Professor Dixon. As a visiting postgraduate student with the University of Manchester, I was able to attend various postgraduate student-related lectures and courses. For my research, I also took the opportunity to learn programming using R for data preparation and analysis and this was made possible with great support and mentorship from Belay Birlie Yimer, a statistician colleague and Anna Beukenhorst, a PhD student with health informatics background. The IT service at the University of Manchester provides programming courses for the staff and students, which I attended to further improve my programming skills in R and Python.

For my research, using the Cloudy with a Chance of Pain dataset, I have the opportunity to examine long-term and short-term day-to-day pain variability in those with musculoskeletal diseases and to examine the driving factors behind the pain variability. First, I descriptively analysed the population-averaged pain severity and other pain symptoms (e.g. pain impact,

mood, fatigue, sleep, waking up tired, morning stiffness, physical activity and wellbeing) over one-month period across different rheumatological conditions such as rheumatoid arthritis (RA), osteoarthritis (OA), spondyloarthropathy (SpA) and fibromyalgia (FM). I also performed similar analysis for participants with comorbid FM in RA, OA and SpA. Secondly, I analysed individual-level pain trajectory in terms of daily pain variability over time and change in pain state over time using different modelling approaches. These analyses are currently in progress and the final work will be disseminated in the form of publications and research presentations at rheumatology conferences. I was able to present the preliminary descriptive analyses of my work at the departmental seminar session in June 2019.

In addition, I am currently conducting a systematic literature review on pain trajectory and pain variability in musculoskeletal diseases and I intend to submit this work in the form of publication. All current and future output from this research will largely form my PhD thesis by publication.

I would like to take this opportunity to acknowledge and to thank Australian Rheumatology Association for this fellowship grant funding, and Arthritis Australia for the opportunity to apply for this fellowship within the National Research Program grants. I would also like to thank Professor Catherine Hill, Dr Samuel Whittle and Dr Rachel Black for their encouragement and support during the fellowship application process. This fellowship has been a rewarding and empowering lifetime experience for me as an early researcher, and I can never thank Professor William Dixon enough for his undivided supervision and mentorship, as well as the *Cloudy* team and colleagues in the department in Manchester.

The role of social media in community perceptions of rheumatoid arthritis drug therapy

Dr Helen Keen

Identifying factors that reduce the risk of bone fractures

Dr Feitong Wu

Delivering rheumatoid arthritis drugs through oral liposomes

Dr Meghna Takelar

Improving outcomes in Rheumatoid Arthritis

Dr Mihir Dilip Wechalekar

Obesity and psoriatic Arthritis

Dr Premarani Sinnathurai

Communicating low back pain research to the consumer

Dr Jenny Setchell

Predicting knee loading using wearable sensors

Dr Amity Campbell and Tara Binnie

Appendix F: Abstracts Presented During PhD Candidature

ABSTRACT NUMBER: 0542

Examining the Long-Term and Short-Term Day-To-Day Pain Variability in Inflammatory and Non-Inflammatory Rheumatic and Musculoskeletal Diseases Using Multilevel and Markov Transition Models: Cloudy with a Chance of Pain, a National U.K. Smartphone Study

Huai Leng Pisaniello[1], Mark Lunt[2], John McBeth[3] and William Dixon[2], [1]University of Adelaide, Kidman Park, South Australia, Australia, [2]Centre for Epidemiology Versus Arthritis, University of Manchester, Manchester, United Kingdom, [3]University of Manchester, Manchester, United Kingdom

Meeting: ACR Convergence 2020

Keywords: Epidemiology, longitudinal studies, Outcome measures, pain

SESSION INFORMATION

Date: Saturday, November 7, 2020

Session Type: Poster Session B

Title: Epidemiology & Public Health Poster II: OA, Osteoporosis, & Other Rheumatic Disease

Session Time: 9:00AM-11:00AM

Background/Purpose: Chronic pain is common in rheumatic and musculoskeletal diseases (RMDs), yet the patterns and the extent of variability over time are poorly understood. Real-time longitudinal capture of pain symptoms using smartphones enables the assessment of temporal patterns of pain, which are important indicators of disease activity. We examined the day-to-day pain variability in inflammatory and non-inflammatory RMDs using *Cloudy With a Chance of Pain*, a national U.K. smartphone study.

Methods: 10,584 study participants (aged ≥18 years; chronic pain for ≥3 months) entered their daily pain using a downloaded smartphone app (on a five-point ordinal scale of 1 – no pain/ 2 – mild pain/ 3 – moderate pain/ 4 – severe pain/ 5 – very severe pain). 2,525 participants diagnosed with single RMD (rheumatoid arthritis – RA, axial spondyloarthropathy – axSpA, osteoarthritis – OA, chronic widespread pain/fibromyalgia – CWP/FM) were included (median symptom entry days: 165 [IQR 82-284]). Long-term and short-term day-to-day pain variability for the first one-month period were analyzed using multilevel and Markov transition models respectively.

Results: From 29,705 daily pain scores (83% female; mean age 48 years; median symptom entry days: 24 [IQR 15-29]), the average pain scores for the first one-month period were higher in participants with axSpA and OA compared with participants with RA (2.74±0.98, 2.61±0.96, and 2.53±0.98 respectively), although participants with CWP/FM had the highest average pain score of 3.06±1.04. In addition, participants with CWP/FM had the highest overall pain level (71.1% reported moderate-very severe pain), followed by participants with axSpA, OA, and RA (57.8%, 52.0%, and 47.9% reported moderate-very severe pain respectively). The long-term change in pain was significantly different between participants, with steeper time-based improvements in pain for participants with higher initial pain scores across all diseases. The day-to-day pain state transitions were unchanged in 50% of days across diseases, although the event of any increase in pain state was

noted in 25% of days (e.g., ≥2-point increase was noted in 4% of days). 53% of those with CWP/FM remained in the 'very severe' pain state with minimal variation.

Conclusion: Participants with CWP/FM had the highest overall pain level followed by participants with axSpA, OA, and RA. These daily pain scores allow the assessment of gradual day-to-day changes through time. Patterns of improvement in those with higher initial pain scores were seen across diseases, perhaps representing regression to the mean. The volatility of changing pain states was comparable across diseases, suggesting no difference in flares. Our future work in identifying patterns of day-to-day pain will focus on analyzing the magnitude of day-to-day change in pain and the constructs of pain volatility.

	RA	axSpA	OA	CWPFM
N	921	216	758	630
Female gender, N(%)	756 (82.1)	143 (66.2)	623 (82.2)	568 (90.2)
Age, mean(SD)	47.40 (12.51)	43.55 (11.02)	55.74 (11.04)	41.25 (10.75)
Pain site - head, N(%)	48 (5.2)	14 (6.5)	24 (3.2)	190 (30.2)
Pain site - face, N(%)	17 (1.8)	2 (0.9)	5 (0.7)	102 (16.2)
Pain site - mouth/jaw, N(%)	140 (15.2)	26 (12.0)	26 (3.4)	170 (27.0)
Pain site - neck/shoulder, N(%)	505 (54.8)	138 (63.9)	284 (37.5)	397 (63.0)
Pain site - back, N(%)	298 (32.4)	191 (88.4)	339 (44.7)	412 (65.4)
Pain site - abdomen, N(%)	46 (5.0)	27 (12.5)	31 (4.1)	195 (31.0)
Pain site - hip, N(%)	384 (41.7)	123 (56.9)	376 (49.6)	363 (57.6)
Pain site - knee, N(%)	568 (61.7)	95 (44.0)	516 (68.1)	376 (59.7)
Pain site - hands, N(%)	716 (77.7)	86 (39.8)	373 (49.2)	346 (54.9)
Pain site - feet, N(%)	574 (62.3)	71 (32.9)	256 (33.8)	301 (47.8)
Pain site - multi-site, N(%)	426 (46.3)	86 (39.8)	128 (16.9)	396 (62.9)
Pain site - all over body, N(%)	107 (11.6)	8 (3.7)	22 (2.9)	331 (52.5)
Analgesia use - none, N(%)	85 (9.2)	11 (5.1)	65 (8.6)	32 (5.1)
Analgesia use - paracetamol, N(%)	410 (44.5)	89 (41.2)	393 (51.8)	302 (47.9)
Analgesia use - NSAIDs, N(%)	580 (63.0)	156 (72.2)	478 (63.1)	336 (53.3)
Analgesia use - simple analgesia, N(%)	277 (30.1)	55 (25.5)	224 (29.6)	221 (35.1)
Analgesia use - weak opioids, N(%)	221 (24.0)	60 (27.8)	186 (24.5)	262 (41.6)
Analgesia use - strong opioids, N(%)	63 (6.8)	21 (9.7)	44 (5.8)	76 (12.1)
Analgesia use - neuropathic pain agent, N(%)	51 (5.5)	26 (12.0)	55 (7.3)	267 (42.4)
Analgesia use - other, N(%)	20 (2.2)	7 (3.2)	49 (6.5)	100 (15.9)

Table 1. Baseline characteristics for 2,525 study participants stratified by rheumatic and musculoskeletal diseases

Figure 1. Slope-intercept plots for the multilevel model stratified by rheumatic and musculoskeletal diseases

Figure 2. Heat map plots for the Markov transition model stratified by rheumatic and musculoskeletal diseases

Disclosure: H. Pisaniello, None; **M. Lunt**, None; **J. McBeth**, None; **W. Dixon**, Google, 1, Bayer, 1.

To cite this abstract in AMA style:

Pisaniello H, Lunt M, McBeth J, Dixon W. Examining the Long-Term and Short-Term Day-To-Day Pain Variability in Inflammatory and Non-Inflammatory Rheumatic and Musculoskeletal Diseases Using Multilevel and Markov Transition Models: Cloudy with a Chance of Pain, a National U.K. Smartphone Study [abstract]. *Arthritis Rheumatol.* 2020; 72 (suppl 10). https://acrabstracts.org/abstract/examining-the-long-term-and-short-term-day-to-day-pain-variability-in-inflammatory-and-non-inflammatory-rheumatic-and-musculoskeletal-diseases-using-multilevel-and-markov-transition-models-cloudy-wit/. Accessed November 29, 2023.

Reference

1. Dickinson E. Pain has an element of blank In *Collected Poems of Emily Dickinson*. Todd, ML, Higginson, TW. New York Avenal Press; 1982.
2. Goldberg DS. The lived experiences of chronic pain. Am J Med. 2012 Aug;125(8):836-7. DOI: 10.1016/j.amjmed.2012.01.032.
3. Gebhart GF. Bonica's Management of Pain; Chapter 1: Intellectual Milestones in Our Understanding and Treatment of Pain. 2018.
4. Geenen R, Overman CL, Christensen R, Asenlof P, Capela S, Huisinga KL, et al. EULAR recommendations for the health professional's approach to pain management in inflammatory arthritis and osteoarthritis. Ann Rheum Dis. 2018 Jun;77(6):797-807. DOI: 10.1136/annrheumdis-2017-212662.
5. Schaible HG, Ebersberger A, Von Banchet GS. Mechanisms of pain in arthritis. Ann N Y Acad Sci. 2002 Jun;966:343-54. DOI: 10.1111/j.1749-6632.2002.tb04234.x.
6. Vladimirova N, Jespersen A, Bartels EM, Christensen AW, Bliddal H, Danneskiold-Samsoe B. Pain Sensitisation in Women with Active Rheumatoid Arthritis: A Comparative Cross-Sectional Study. Arthritis. 2015;2015:434109. DOI: 10.1155/2015/434109.
7. Zhang A, Lee YC. Mechanisms for Joint Pain in Rheumatoid Arthritis (RA): from Cytokines to Central Sensitization. Curr Osteoporos Rep. 2018 Oct;16(5):603-10. DOI: 10.1007/s11914-018-0473-5.
8. Aletaha D, Smolen JS. Diagnosis and Management of Rheumatoid Arthritis: A Review. JAMA. 2018 Oct 2;320(13):1360-72. DOI: 10.1001/jama.2018.13103.
9. Smolen JS, Aletaha D, Barton A, Burmester GR, Emery P, Firestein GS, et al. Rheumatoid arthritis. Nat Rev Dis Primers. 2018 Feb 8;4:18001. DOI: 10.1038/nrdp.2018.1.
10. Sparks JA. Rheumatoid Arthritis. Ann Intern Med. 2019 Jan 1;170(1):ITC1-ITC16. DOI: 10.7326/AITC201901010.
11. (AIHW) AIoHaW. Chronic musculoskeletal conditions: Rheumatoid arthritis. [updated 9 February 2023]. Available from: www.aihw.gov.au/reports/chronic-musculoskeletal-conditions/musculoskeletal-conditions/contents/arthritis/rheumatoid-arthritis
12. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature. 2014 Feb 20;506(7488):376-81. DOI: 10.1038/nature12873.
13. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet. 2012 Jan 29;44(3):291-6. DOI: 10.1038/ng.1076.
14. Scott IC, Seegobin SD, Steer S, Tan R, Forabosco P, Hinks A, et al. Predicting the risk of rheumatoid arthritis and its age of onset through modelling genetic risk variants with smoking. PLoS Genet. 2013;9(9):e1003808. DOI: 10.1371/journal.pgen.1003808.
15. Silman AJ, Newman J, MacGregor AJ. Cigarette smoking increases the risk of rheumatoid arthritis. Results from a nationwide study of disease-discordant twins. Arthritis Rheum. 1996 May;39(5):732-5. DOI: 10.1002/art.1780390504.
16. Sugiyama D, Nishimura K, Tamaki K, Tsuji G, Nakazawa T, Morinobu A, et al. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. Ann Rheum Dis. 2010 Jan;69(1):70-81. DOI: 10.1136/ard.2008.096487.

17. Viatte S, Barton A. Genetics of rheumatoid arthritis susceptibility, severity, and treatment response. Semin Immunopathol. 2017 Jun;39(4):395-408. DOI: 10.1007/s00281-017-0630-4.
18. Ospelt C, Gay S, Klein K. Epigenetics in the pathogenesis of RA. Semin Immunopathol. 2017 Jun;39(4):409-19. DOI: 10.1007/s00281-017-0621-5.
19. Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum. 2004 Feb;50(2):380-6. DOI: 10.1002/art.20018.
20. Favalli EG, Becciolini A, Biggioggero M. Structural integrity versus radiographic progression in rheumatoid arthritis. RMD Open. 2015;1(Suppl 1):e000064. DOI: 10.1136/rmdopen-2015-000064.
21. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJ, Hazes JM, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. Arthritis Rheum. 2005 Nov;52(11):3381-90. DOI: 10.1002/art.21405.
22. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010 Sep;62(9):2569-81. DOI: 10.1002/art.27584.
23. Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A, et al. Distinctions between diagnostic and classification criteria? Arthritis Care Res (Hoboken). 2015 Jul;67(7):891-7. DOI: 10.1002/acr.22583.
24. Radner H, Neogi T, Smolen JS, Aletaha D. Performance of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: a systematic literature review. Ann Rheum Dis. 2014 Jan;73(1):114-23. DOI: 10.1136/annrheumdis-2013-203284.
25. Emery P, Breedveld FC, Dougados M, Kalden JR, Schiff MH, Smolen JS. Early referral recommendation for newly diagnosed rheumatoid arthritis: evidence based development of a clinical guide. Ann Rheum Dis. 2002 Apr;61(4):290-7. DOI: 10.1136/ard.61.4.290.
26. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther. 2005;7(4):R796-806. DOI: 10.1186/ar1740.
27. van der Heijde DM, van Riel PL, van Leeuwen MA, van 't Hof MA, van Rijswijk MH, van de Putte LB. Prognostic factors for radiographic damage and physical disability in early rheumatoid arthritis. A prospective follow-up study of 147 patients. Br J Rheumatol. 1992 Aug;31(8):519-25. DOI: 10.1093/rheumatology/31.8.519.
28. Aletaha D, Alasti F, Smolen JS. Optimisation of a treat-to-target approach in rheumatoid arthritis: strategies for the 3-month time point. Ann Rheum Dis. 2016 Aug;75(8):1479-85. DOI: 10.1136/annrheumdis-2015-208324.
29. Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis. 2016 Jan;75(1):3-15. DOI: 10.1136/annrheumdis-2015-207524.
30. Brandt KD, Radin EL, Dieppe PA, van de Putte L. Yet more evidence that osteoarthritis is not a cartilage disease. Ann Rheum Dis. 2006 Oct;65(10):1261-4. DOI: 10.1136/ard.2006.058347.
31. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. Lancet. 2019 Apr 27;393(10182):1745-59. DOI: 10.1016/S0140-6736(19)30417-9.

32. Martel-Pelletier J, Barr AJ, Cicuttini FM, Conaghan PG, Cooper C, Goldring MB, et al. Osteoarthritis. Nat Rev Dis Primers. 2016 Oct 13;2:16072. DOI: 10.1038/nrdp.2016.72.
33. Mathiessen A, Conaghan PG. Synovitis in osteoarthritis: current understanding with therapeutic implications. Arthritis Res Ther. 2017 Feb 2;19(1):18. DOI: 10.1186/s13075-017-1229-9.
34. Disease GBD, Injury I, Prevalence C. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016 Oct 8;388(10053):1545-602. DOI: 10.1016/S0140-6736(16)31678-6.
35. Prieto-Alhambra D, Judge A, Javaid MK, Cooper C, Diez-Perez A, Arden NK. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: influences of age, gender and osteoarthritis affecting other joints. Ann Rheum Dis. 2014 Sep;73(9):1659-64. DOI: 10.1136/annrheumdis-2013-203355.
36. Turkiewicz A, Petersson IF, Bjork J, Hawker G, Dahlberg LE, Lohmander LS, et al. Current and future impact of osteoarthritis on health care: a population-based study with projections to year 2032. Osteoarthritis Cartilage. 2014 Nov;22(11):1826-32. DOI: 10.1016/j.joca.2014.07.015.
37. (AIHW) AIoHaW. Chronic musculoskeletal conditions: Osteoarthritis. [updated 9 February 2023]. Available from: https://www.aihw.gov.au/reports/chronic-musculoskeletal-conditions/musculoskeletal-conditions/contents/arthritis/osteoarthritis
38. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Rheum Dis Clin North Am. 2008 Aug;34(3):515-29. DOI: 10.1016/j.rdc.2008.05.007.
39. Driban JB, Hootman JM, Sitler MR, Harris KP, Cattano NM. Is Participation in Certain Sports Associated With Knee Osteoarthritis? A Systematic Review. J Athl Train. 2017 Jun 2;52(6):497-506. DOI: 10.4085/1062-6050-50.2.08.
40. Ezzat AM, Li LC. Occupational physical loading tasks and knee osteoarthritis: a review of the evidence. Physiother Can. 2014 Winter;66(1):91-107. DOI: 10.3138/ptc.2012-45BC.
41. Harris EC, Coggon D. HIP osteoarthritis and work. Best Pract Res Clin Rheumatol. 2015 Jun;29(3):462-82. DOI: 10.1016/j.berh.2015.04.015.
42. Vigdorchik JM, Nepple JJ, Eftekhary N, Leunig M, Clohisy JC. What Is the Association of Elite Sporting Activities With the Development of Hip Osteoarthritis? Am J Sports Med. 2017 Mar;45(4):961-4. DOI: 10.1177/0363546516656359.
43. Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. Osteoarthritis Cartilage. 2015 Apr;23(4):507-15. DOI: 10.1016/j.joca.2014.11.019.
44. Block JA. Osteoarthritis: OA guidelines: improving care or merely codifying practice? Nat Rev Rheumatol. 2014 Jun;10(6):324-6. DOI: 10.1038/nrrheum.2014.61.
45. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: The chronic osteoarthritis management initiative of the U.S. bone and joint initiative. Semin Arthritis Rheum. 2014 Jun;43(6):701-12. DOI: 10.1016/j.semarthrit.2013.11.012.
46. Robinson PC, van der Linden S, Khan MA, Taylor WJ. Axial spondyloarthritis: concept, construct, classification and implications for therapy. Nat Rev Rheumatol. 2021 Feb;17(2):109-18. DOI: 10.1038/s41584-020-00552-4.

47. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum. 1984 Apr;27(4):361-8. DOI: 10.1002/art.1780270401.
48. Sarzi-Puttini P, Giorgi V, Marotto D, Atzeni F. Fibromyalgia: an update on clinical characteristics, aetiopathogenesis and treatment. Nat Rev Rheumatol. 2020 Nov;16(11):645-60. DOI: 10.1038/s41584-020-00506-w.
49. Queiroz LP. Worldwide epidemiology of fibromyalgia. Curr Pain Headache Rep. 2013 Aug;17(8):356. DOI: 10.1007/s11916-013-0356-5.
50. White KP, Speechley M, Harth M, Ostbye T. The London Fibromyalgia Epidemiology Study: the prevalence of fibromyalgia syndrome in London, Ontario. J Rheumatol. 1999 Jul;26(7):1570-6.
51. Hauser W, Sarzi-Puttini P, Fitzcharles MA. Fibromyalgia syndrome: under-, over- and misdiagnosis. Clin Exp Rheumatol. 2019 Jan-Feb;37 Suppl 116(1):90-7.
52. Giusti EM, Castelnuovo G, Molinari E. Differences in Multidisciplinary and Interdisciplinary Treatment Programs for Fibromyalgia: A Mapping Review. Pain Res Manag. 2017;2017:7261468. DOI: 10.1155/2017/7261468.
53. Haase L, Stewart JL, Youssef B, May AC, Isakovic S, Simmons AN, et al. When the brain does not adequately feel the body: Links between low resilience and interoception. Biol Psychol. 2016 Jan;113:37-45. DOI: 10.1016/j.biopsycho.2015.11.004.
54. Yunus MB. Central sensitivity syndromes: a new paradigm and group nosology for fibromyalgia and overlapping conditions, and the related issue of disease versus illness. Semin Arthritis Rheum. 2008 Jun;37(6):339-52. DOI: 10.1016/j.semarthrit.2007.09.003.
55. Rocha e Silva M. A brief survey of the history of inflammation. Agents Actions. 1978 Jan;8(1-2):45-9. DOI: 10.1007/BF01972401.
56. Scott A, Khan KM, Cook JL, Duronio V. What is "inflammation"? Are we ready to move beyond Celsus? Br J Sports Med. 2004 Jun;38(3):248-9. DOI: 10.1136/bjsm.2003.011221.
57. Granger DN, Senchenkova E. Chapter 2 Historical Perspective In Inflammation and the Microcirculation. In *Integrated Systems Physiology-From Cell to Function*. San Rafael (CA)2010.
58. Bayliss WM. On the origin from the spinal cord of the vaso-dilator fibres of the hind-limb, and on the nature of these fibres. J Physiol. 1901 Feb 28;26(3-4):173-209. DOI: 10.1113/jphysiol.1901.sp000831.
59. Mach DB, Rogers SD, Sabino MC, Luger NM, Schwei MJ, Pomonis JD, et al. Origins of skeletal pain: sensory and sympathetic innervation of the mouse femur. Neuroscience. 2002;113(1):155-66. DOI: 10.1016/s0306-4522(02)00165-3.
60. McDougall JJ. Arthritis and pain. Neurogenic origin of joint pain. Arthritis Res Ther. 2006;8(6):220. DOI: 10.1186/ar2069.
61. McDougall JJ, Bray RC, Sharkey KA. Morphological and immunohistochemical examination of nerves in normal and injured collateral ligaments of rat, rabbit, and human knee joints. Anat Rec. 1997 May;248(1):29-39. DOI: 10.1002/(SICI)1097-0185(199705)248:1<29::AID-AR4>3.0.CO;2-A.
62. Samuel EP. The autonomic and somatic innervation of the articular capsule. Anat Rec. 1952 May;113(1):53-70. DOI: 10.1002/ar.1091130104.
63. Grassel S, Muschter D. Peripheral Nerve Fibers and Their Neurotransmitters in Osteoarthritis Pathology. Int J Mol Sci. 2017 Apr 28;18(5). DOI: 10.3390/ijms18050931.
64. Fu K, Robbins SR, McDougall JJ. Osteoarthritis: the genesis of pain. Rheumatology (Oxford). 2018 May 1;57(suppl_4):iv43-iv50. DOI: 10.1093/rheumatology/kex419.

65. Walsh DA, McWilliams DF, Turley MJ, Dixon MR, Franses RE, Mapp PI, et al. Angiogenesis and nerve growth factor at the osteochondral junction in rheumatoid arthritis and osteoarthritis. Rheumatology (Oxford). 2010 Oct;49(10):1852-61. DOI: 10.1093/rheumatology/keq188.
66. Arnalich F, de Miguel E, Perez-Ayala C, Martinez M, Vazquez JJ, Gijon-Banos J, et al. Neuropeptides and interleukin-6 in human joint inflammation relationship between intraarticular substance P and interleukin-6 concentrations. Neurosci Lett. 1994 Apr 11;170(2):251-4. DOI: 10.1016/0304-3940(94)90331-x.
67. Hernanz A, De Miguel E, Romera N, Perez-Ayala C, Gijon J, Arnalich F. Calcitonin gene-related peptide II, substance P and vasoactive intestinal peptide in plasma and synovial fluid from patients with inflammatory joint disease. Br J Rheumatol. 1993 Jan;32(1):31-5. DOI: 10.1093/rheumatology/32.1.31.
68. McWilliams DF, Walsh DA. Pain mechanisms in rheumatoid arthritis. Clin Exp Rheumatol. 2017 Sep-Oct;35 Suppl 107(5):94-101.
69. Nwosu LN, Mapp PI, Chapman V, Walsh DA. Blocking the tropomyosin receptor kinase A (TrkA) receptor inhibits pain behaviour in two rat models of osteoarthritis. Ann Rheum Dis. 2016 Jun;75(6):1246-54. DOI: 10.1136/annrheumdis-2014-207203.
70. Ogbonna AC, Clark AK, Malcangio M. Development of monosodium acetate-induced osteoarthritis and inflammatory pain in ageing mice. Age (Dordr). 2015 Jun;37(3):9792. DOI: 10.1007/s11357-015-9792-y.
71. Nishioku T, Yamauchi A, Takata F, Watanabe T, Furusho K, Shuto H, et al. Disruption of the blood-brain barrier in collagen-induced arthritic mice. Neurosci Lett. 2010 Oct 4;482(3):208-11. DOI: 10.1016/j.neulet.2010.07.035.
72. Nieto FR, Clark AK, Grist J, Hathway GJ, Chapman V, Malcangio M. Neuron-immune mechanisms contribute to pain in early stages of arthritis. J Neuroinflammation. 2016 Apr 29;13(1):96. DOI: 10.1186/s12974-016-0556-0.
73. Suter MR, Wen YR, Decosterd I, Ji RR. Do glial cells control pain? Neuron Glia Biol. 2007 Aug;3(3):255-68. DOI: 10.1017/S1740925X08000100.
74. Schaible HG, von Banchet GS, Boettger MK, Brauer R, Gajda M, Richter F, et al. The role of proinflammatory cytokines in the generation and maintenance of joint pain. Ann N Y Acad Sci. 2010 Apr;1193:60-9. DOI: 10.1111/j.1749-6632.2009.05301.x.
75. Nieto FR, Clark AK, Grist J, Chapman V, Malcangio M. Calcitonin gene-related peptide-expressing sensory neurons and spinal microglial reactivity contribute to pain states in collagen-induced arthritis. Arthritis Rheumatol. 2015 Jun;67(6):1668-77. DOI: 10.1002/art.39082.
76. Salaffi F, Giacobazzi G, Di Carlo M. Chronic Pain in Inflammatory Arthritis: Mechanisms, Metrology, and Emerging Targets-A Focus on the JAK-STAT Pathway. Pain Res Manag. 2018;2018:8564215. DOI: 10.1155/2018/8564215.
77. Pain terms: a list with definitions and notes on usage. Recommended by the IASP Subcommittee on Taxonomy. Pain. 1979 Jun;6(3):249. Cited in: Medline.
78. Raja SN, Carr DB, Cohen M, Finnerup NB, Flor H, Gibson S, et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. Pain. 2020 Sep 1;161(9):1976-82. DOI: 10.1097/j.pain.0000000000001939. Cited in: Medline.
79. Altawil R, Saevarsdottir S, Wedren S, Alfredsson L, Klareskog L, Lampa J. Remaining Pain in Early Rheumatoid Arthritis Patients Treated With Methotrexate. Arthritis Care Res (Hoboken). 2016 Aug;68(8):1061-8. DOI: 10.1002/acr.22790.
80. Lee YC, Cui J, Lu B, Frits ML, Iannaccone CK, Shadick NA, et al. Pain persists in DAS28 rheumatoid arthritis remission but not in ACR/EULAR remission: a

longitudinal observational study. Arthritis Res Ther. 2011 Jun 8;13(3):R83. DOI: 10.1186/ar3353.
81. McWilliams DF, Walsh DA. Factors predicting pain and early discontinuation of tumour necrosis factor-alpha-inhibitors in people with rheumatoid arthritis: results from the British society for rheumatology biologics register. BMC Musculoskelet Disord. 2016 Aug 12;17:337. DOI: 10.1186/s12891-016-1192-7.
82. Suokas AK, Walsh DA, McWilliams DF, Condon L, Moreton B, Wylde V, et al. Quantitative sensory testing in painful osteoarthritis: a systematic review and meta-analysis. Osteoarthritis Cartilage. 2012 Oct;20(10):1075-85. DOI: 10.1016/j.joca.2012.06.009.
83. Friden C, Thoors U, Glenmark B, Kosek E, Nordmark B, Lundberg IE, et al. Higher pain sensitivity and lower muscle strength in postmenonpausal women with early rheumatoid arthritis compared with age-matched healthy women--a pilot study. Disabil Rehabil. 2013 Aug;35(16):1350-6. DOI: 10.3109/09638288.2012.731469.
84. Leffler AS, Kosek E, Lerndal T, Nordmark B, Hansson P. Somatosensory perception and function of diffuse noxious inhibitory controls (DNIC) in patients suffering from rheumatoid arthritis. Eur J Pain. 2002;6(2):161-76. DOI: 10.1053/eujp.2001.0313.
85. Meeus M, Vervisch S, De Clerck LS, Moorkens G, Hans G, Nijs J. Central sensitization in patients with rheumatoid arthritis: a systematic literature review. Semin Arthritis Rheum. 2012 Feb;41(4):556-67. DOI: 10.1016/j.semarthrit.2011.08.001.
86. Joharatnam N, McWilliams DF, Wilson D, Wheeler M, Pande I, Walsh DA. A cross-sectional study of pain sensitivity, disease-activity assessment, mental health, and fibromyalgia status in rheumatoid arthritis. Arthritis Res Ther. 2015 Jan 20;17(1):11. DOI: 10.1186/s13075-015-0525-5.
87. Lee YC, Chibnik LB, Lu B, Wasan AD, Edwards RR, Fossel AH, et al. The relationship between disease activity, sleep, psychiatric distress and pain sensitivity in rheumatoid arthritis: a cross-sectional study. Arthritis Res Ther. 2009;11(5):R160. DOI: 10.1186/ar2842.
88. Lee YC, Lu B, Edwards RR, Wasan AD, Nassikas NJ, Clauw DJ, et al. The role of sleep problems in central pain processing in rheumatoid arthritis. Arthritis Rheum. 2013 Jan;65(1):59-68. DOI: 10.1002/art.37733.
89. Sim MK, Kim DY, Yoon J, Park DH, Kim YG. Assessment of peripheral neuropathy in patients with rheumatoid arthritis who complain of neurologic symptoms. Ann Rehabil Med. 2014 Apr;38(2):249-55. DOI: 10.5535/arm.2014.38.2.249.
90. Ahmed S, Magan T, Vargas M, Harrison A, Sofat N. Use of the painDETECT tool in rheumatoid arthritis suggests neuropathic and sensitization components in pain reporting. J Pain Res. 2014;7:579-88. DOI: 10.2147/JPR.S69011.
91. Freynhagen R, Baron R, Gockel U, Tolle TR. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Curr Med Res Opin. 2006 Oct;22(10):1911-20. DOI: 10.1185/030079906X132488.
92. Gauffin J, Hankama T, Kautiainen H, Hannonen P, Haanpaa M. Neuropathic pain and use of PainDETECT in patients with fibromyalgia: a cohort study. BMC Neurol. 2013 Feb 14;13:21. DOI: 10.1186/1471-2377-13-21.
93. Hochman JR, Davis AM, Elkayam J, Gagliese L, Hawker GA. Neuropathic pain symptoms on the modified painDETECT correlate with signs of central sensitization in knee osteoarthritis. Osteoarthritis Cartilage. 2013 Sep;21(9):1236-42. DOI: 10.1016/j.joca.2013.06.023.

94. Morso L, Kent PM, Albert HB. Are self-reported pain characteristics, classified using the PainDETECT questionnaire, predictive of outcome in people with low back pain and associated leg pain? Clin J Pain. 2011 Jul-Aug;27(6):535-41. DOI: 10.1097/AJP.0b013e318208c941.
95. Banic B, Petersen-Felix S, Andersen OK, Radanov BP, Villiger PM, Arendt-Nielsen L, et al. Evidence for spinal cord hypersensitivity in chronic pain after whiplash injury and in fibromyalgia. Pain. 2004 Jan;107(1-2):7-15. DOI: 10.1016/j.pain.2003.05.001.
96. Meeus M, Nijs J. Central sensitization: a biopsychosocial explanation for chronic widespread pain in patients with fibromyalgia and chronic fatigue syndrome. Clin Rheumatol. 2007 Apr;26(4):465-73. DOI: 10.1007/s10067-006-0433-9.
97. Meeus M, Nijs J, Huybrechts S, Truijen S. Evidence for generalized hyperalgesia in chronic fatigue syndrome: a case control study. Clin Rheumatol. 2010 Apr;29(4):393-8. DOI: 10.1007/s10067-009-1339-0.
98. Nijs J, Meeus M, Van Oosterwijck J, Ickmans K, Moorkens G, Hans G, et al. In the mind or in the brain? Scientific evidence for central sensitisation in chronic fatigue syndrome. Eur J Clin Invest. 2012 Feb;42(2):203-12. DOI: 10.1111/j.1365-2362.2011.02575.x.
99. Lee YC, Lu B, Boire G, Haraoui BP, Hitchon CA, Pope JE, et al. Incidence and predictors of secondary fibromyalgia in an early arthritis cohort. Ann Rheum Dis. 2013 Jun;72(6):949-54. DOI: 10.1136/annrheumdis-2012-201506.
100. Wolfe F, Hauser W, Hassett AL, Katz RS, Walitt BT. The development of fibromyalgia--I: examination of rates and predictors in patients with rheumatoid arthritis (RA). Pain. 2011 Feb;152(2):291-9. DOI: 10.1016/j.pain.2010.09.027.
101. Coury F, Rossat A, Tebib A, Letroublon MC, Gagnard A, Fantino B, et al. Rheumatoid arthritis and fibromyalgia: a frequent unrelated association complicating disease management. J Rheumatol. 2009 Jan;36(1):58-62. DOI: 10.3899/jrheum.080366.
102. Pollard LC, Ibrahim F, Choy EH, Scott DL. Pain thresholds in rheumatoid arthritis: the effect of tender point counts and disease duration. J Rheumatol. 2012 Jan;39(1):28-31. DOI: 10.3899/jrheum.110668.
103. Pollard LC, Kingsley GH, Choy EH, Scott DL. Fibromyalgic rheumatoid arthritis and disease assessment. Rheumatology (Oxford). 2010 May;49(5):924-8. DOI: 10.1093/rheumatology/kep458.
104. Ranzolin A, Brenol JC, Bredemeier M, Guarienti J, Rizzatti M, Feldman D, et al. Association of concomitant fibromyalgia with worse disease activity score in 28 joints, health assessment questionnaire, and short form 36 scores in patients with rheumatoid arthritis. Arthritis Rheum. 2009 Jun 15;61(6):794-800. DOI: 10.1002/art.24430.
105. Wolfe F, Cathey MA, Kleinheksel SM. Fibrositis (Fibromyalgia) in rheumatoid arthritis. J Rheumatol. 1984 Dec;11(6):814-8.
106. Wolfe F, Cathey MA, Kleinheksel SM, Amos SP, Hoffman RG, Young DY, et al. Psychological status in primary fibrositis and fibrositis associated with rheumatoid arthritis. J Rheumatol. 1984 Aug;11(4):500-6.
107. McWilliams DF, Marshall M, Jayakumar K, Doherty S, Doherty M, Zhang W, et al. Erosive and osteoarthritic structural progression in early rheumatoid arthritis. Rheumatology (Oxford). 2016 Aug;55(8):1477-88. DOI: 10.1093/rheumatology/kew197.
108. McWilliams DF, Zhang W, Mansell JS, Kiely PD, Young A, Walsh DA. Predictors of change in bodily pain in early rheumatoid arthritis: an inception cohort study. Arthritis Care Res (Hoboken). 2012 Oct;64(10):1505-13. DOI: 10.1002/acr.21723.

109. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. Spine (Phila Pa 1976). 2005 Jun 1;30(11):1331-4. DOI: 10.1097/01.brs.0000164099.92112.29.
110. Huskisson EC. Measurement of pain. Lancet. 1974 Nov 9;2(7889):1127-31. DOI: 10.1016/s0140-6736(74)90884-8.
111. Jensen MP, McFarland CA. Increasing the reliability and validity of pain intensity measurement in chronic pain patients. Pain. 1993 Nov;55(2):195-203. DOI: 10.1016/0304-3959(93)90148-I.
112. Keele KD. The pain chart. Lancet. 1948 Jul 3;2(6514):6-8. DOI: 10.1016/s0140-6736(48)91787-5.
113. Scott J, Huskisson EC. Graphic representation of pain. Pain. 1976 Jun;2(2):175-84.
114. Anderson DL. Development of an instrument to measure pain in rheumatoid arthritis: Rheumatoid Arthritis Pain Scale (RAPS). Arthritis Rheum. 2001 Aug;45(4):317-23. DOI: 10.1002/1529-0131(200108)45:4<317::AID-ART343>3.0.CO;2-X.
115. Bellamy N. Pain assessment in osteoarthritis: experience with the WOMAC osteoarthritis index. Semin Arthritis Rheum. 1989 May;18(4 Suppl 2):14-7. DOI: 10.1016/0049-0172(89)90010-3.
116. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988 Dec;15(12):1833-40.
117. Boureau F, Luu M, Doubrere JF, Gay C. [Construction of a questionnaire for the self-evaluation of pain using a list of qualifiers. Comparison with Melzack's McGill Pain Questionnaire]. Therapie. 1984 Mar-Apr;39(2):119-29.
118. Hawker GA, Davis AM, French MR, Cibere J, Jordan JM, March L, et al. Development and preliminary psychometric testing of a new OA pain measure--an OARSI/OMERACT initiative. Osteoarthritis Cartilage. 2008 Apr;16(4):409-14. DOI: 10.1016/j.joca.2007.12.015.
119. Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). Pain. 1985 Dec;23(4):345-56. DOI: 10.1016/0304-3959(85)90004-1.
120. Leplege A, Ecosse E, Verdier A, Perneger TV. The French SF-36 Health Survey: translation, cultural adaptation and preliminary psychometric evaluation. J Clin Epidemiol. 1998 Nov;51(11):1013-23. DOI: 10.1016/s0895-4356(98)00093-6.
121. Meenan RF, Gertman PM, Mason JH. Measuring health status in arthritis. The arthritis impact measurement scales. Arthritis Rheum. 1980 Feb;23(2):146-52. DOI: 10.1002/art.1780230203.
122. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. Pain. 1975 Sep;1(3):277-99. DOI: 10.1016/0304-3959(75)90044-5.
123. Melzack R. The McGill Pain Questionnaire. In Pain Measurement and Assessment. New York: Raven Press; 1983.
124. Melzack R. The short-form McGill Pain Questionnaire. Pain. 1987 Aug;30(2):191-7. DOI: 10.1016/0304-3959(87)91074-8.
125. Uhlig T, Moe RH, Kvien TK. The burden of disease in rheumatoid arthritis. Pharmacoeconomics. 2014 Sep;32(9):841-51. DOI: 10.1007/s40273-014-0174-6.
126. Favill D, Rennick C. A case of family periodic paralysis. Archives of Neurology and Psychiatry. 1924;11:674-9.
127. Csikszentmihalyi M, Larson R. Validity and reliability of the Experience-Sampling Method. J Nerv Ment Dis. 1987 Sep;175(9):526-36. DOI: 10.1097/00005053-198709000-00004.

128. Csikszentmihalyi M. Flow: the psychology of optimal experience. ed. . sHP e, editor. New York: Harper Collins; 1991.
129. Delespaul PA, deVries MW. The daily life of ambulatory chronic mental patients. J Nerv Ment Dis. 1987 Sep;175(9):537-44. DOI: 10.1097/00005053-198709000-00005.
130. Stone AA, Neale JM. Effects of severe daily events on mood. J Pers Soc Psychol. 1984 Jan;46(1):137-44. DOI: 10.1037//0022-3514.46.1.137.
131. Stone AA, Reed BR, Neale JM. Changes in daily event frequency precede episodes of physical symptoms. J Human Stress. 1987 Summer;13(2):70-4. DOI: 10.1080/0097840X.1987.9936797.
132. Bassett DR, Jr. Validity and reliability issues in objective monitoring of physical activity. Res Q Exerc Sport. 2000 Jun;71(2 Suppl):S30-6.
133. Montoye H, Kemper H, Saris W, RA W. Measuring physical activity and energy expenditure Champaign, IL: Human Kinetics; 1996.
134. Stone AA, S S. Ecological momentary assessment (EMA) in behavioural medicine. Annals of Behavioural Medicine. 1994;16:199-202.
135. Ross J, Schreiber I, Shiffman S, Stone AA. Chapter 1 Historical Roots and Rationale of Ecological Momentary Assessment (EMA). In *The Science of Real-Time Data Capture: Self-Reports in Health Research*: Oxford University Press Inc; 2007.
136. B.L. F. Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. Cognition Emotion. 2000;14(4):577-606.
137. Redelmeier DA, Kahneman D. Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. Pain. 1996 Jul;66(1):3-8. DOI: 10.1016/0304-3959(96)02994-6.
138. Redelmeier DA, Katz J, Kahneman D. Memories of colonoscopy: a randomized trial. Pain. 2003 Jul;104(1-2):187-94. DOI: 10.1016/s0304-3959(03)00003-4.
139. Varey C., Kahneman D. Experiences extended across time: evaluation of moments and episodes. J Behavioural Decision Making. 1992;5:169-85.
140. Stone AA, Broderick JE, Kaell AT, DelesPaul PA, Porter LE. Does the peak-end phenomenon observed in laboratory pain studies apply to real-world pain in rheumatoid arthritics? J Pain. 2000 Autumn;1(3):212-7. DOI: 10.1054/jpai.2000.7568.
141. Istepanian RSH, Pattichis CS, S L. Ubiquitous M-Health Systems and the Convergence Towards 4G Mobile Technologies. In M-Health: Emerging Mobile Health System. Boston, MA: Springer; 2006.
142. Statista. Number of mobile phone users worldwide from 2015 to 2020 (in billions). 2018 [12 August 2018]. Available from: https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/
143. Boulos MN, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomed Eng Online. 2011 Apr 5;10:24. DOI: 10.1186/1475-925X-10-24.
144. U V. Mobile health: four emerging themes of research. Decision Support Systems. 2014;66:20-35.
145. Asada HH, Shaltis P, Reisner A, Rhee S, Hutchinson RC. Mobile monitoring with wearable photoplethysmographic biosensors. IEEE Eng Med Biol Mag. 2003 May-Jun;22(3):28-40. DOI: 10.1109/memb.2003.1213624.
146. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth. 2015 Nov 4;3(4):e101. DOI: 10.2196/mhealth.4924.

147. Luo D, Wang P, Lu F, Elias J, Sparks JA, Lee YC. Mobile Apps for Individuals With Rheumatoid Arthritis: A Systematic Review. J Clin Rheumatol. 2019 Apr;25(3):133-41. DOI: 10.1097/RHU.0000000000000800.
148. Hektner JM, Schmidt JA, Csikszentmihalyi M. Experience Sampling Method: Measuring the quality of everyday life. Thousand Oaks, California: SAGE; 2007.
149. Talbot WJ, Schafer JL, TA W. Multilevel Models for Intensive Longitudinal Data. In *Models for Intensive Longitudinal Data* USA-OSO: Oxford University Press; 2006.
150. Garcia-Palacios A, Herrero R, Belmonte MA, Castilla D, Guixeres J, Molinari G, et al. Ecological momentary assessment for chronic pain in fibromyalgia using a smartphone: a randomized crossover study. Eur J Pain. 2014 Jul;18(6):862-72. DOI: 10.1002/j.1532-2149.2013.00425.x.
151. Beukenhorst AL, Sergeant JC, Little MA, McBeth J, Dixon WG. Consumer Smartwatches for Collecting Self-Report and Sensor Data: App Design and Engagement. Stud Health Technol Inform. 2018;247:291-5.
152. 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 May 21;17:220. DOI: 10.1186/s12891-016-1071-2.
153. Axen I, Leboeuf-Yde C. Trajectories of low back pain. Best Pract Res Clin Rheumatol. 2013 Oct;27(5):601-12. DOI: 10.1016/j.berh.2013.10.004.
154. El Miedany Y. e-Rheumatology: are we ready? Clin Rheumatol. 2015 May;34(5):831-7. DOI: 10.1007/s10067-015-2897-y.
155. El Miedany Y, El Gaafary M, Youssef S, Bahlas S, Almedany S, Ahmed I, et al. Toward Electronic Health Recording: Evaluation of Electronic Patient-reported Outcome Measures System for Remote Monitoring of Early Rheumatoid Arthritis. J Rheumatol. 2016 Dec;43(12):2106-12. DOI: 10.3899/jrheum.151421.
156. Nishiguchi S, Ito H, Yamada M, Yoshitomi H, Furu M, Ito T, et al. Self-assessment of Rheumatoid Arthritis Disease Activity Using a Smartphone Application. Development and 3-month Feasibility Study. Methods Inf Med. 2016;55(1):65-9. DOI: 10.3414/ME14-01-0106.
157. Dixon WG, Michaud K. Using technology to support clinical care and research in rheumatoid arthritis. Curr Opin Rheumatol. 2018 May;30(3):276-81. DOI: 10.1097/BOR.0000000000000485.
158. Callahan D. The WHO definition of 'health'. Stud Hastings Cent. 1973;1(3):77-88.
159. Do. H. What is health? 2022 [cited 2022]. Available from: https://www.aihw.gov.au/reports/australias-health/what-is-health
160. Rumsfeld JS. Health status and clinical practice: when will they meet? Circulation. 2002 Jul 2;106(1):5-7. DOI: 10.1161/01.cir.0000020805.31531.48.
161. Heiberg T, Finset A, Uhlig T, Kvien TK. Seven year changes in health status and priorities for improvement of health in patients with rheumatoid arthritis. Ann Rheum Dis. 2005 Feb;64(2):191-5. DOI: 10.1136/ard.2004.022699.
162. Heiberg T, Kvien TK. Preferences for improved health examined in 1,024 patients with rheumatoid arthritis: pain has highest priority. Arthritis Rheum. 2002 Aug;47(4):391-7. DOI: 10.1002/art.10515.
163. Minnock P, FitzGerald O, Bresnihan B. Women with established rheumatoid arthritis perceive pain as the predominant impairment of health status. Rheumatology (Oxford). 2003 Aug;42(8):995-1000. DOI: 10.1093/rheumatology/keg281.
164. Sanderson T, Morris M, Calnan M, Richards P, Hewlett S. Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. Arthritis Care Res (Hoboken). 2010 May;62(5):647-56. DOI: 10.1002/acr.20151.

165. Barton JL, Imboden J, Graf J, Glidden D, Yelin EH, Schillinger D. Patient-physician discordance in assessments of global disease severity in rheumatoid arthritis. Arthritis Care Res (Hoboken). 2010 Jun;62(6):857-64. DOI: 10.1002/acr.20132.
166. Courvoisier DS, Agoritsas T, Glauser J, Michaud K, Wolfe F, Cantoni E, et al. Pain as an important predictor of psychosocial health in patients with rheumatoid arthritis. Arthritis Care Res (Hoboken). 2012 Feb;64(2):190-6. DOI: 10.1002/acr.20652.
167. Katchamart W, Narongroeknawin P, Chanapai W, Thaweeratthakul P. Health-related quality of life in patients with rheumatoid arthritis. BMC Rheumatol. 2019;3:34. DOI: 10.1186/s41927-019-0080-9.
168. Malm K, Bergman S, Andersson ML, Bremander A, Larsson I. Quality of life in patients with established rheumatoid arthritis: A phenomenographic study. SAGE Open Med. 2017;5:2050312117713647. DOI: 10.1177/2050312117713647.
169. Martinec R, Pinjatela R, Balen D. Quality of Life in Patients with Rheumatoid Arthritis - a Preliminary Study. Acta Clin Croat. 2019 Mar;58(1):157-66. DOI: 10.20471/acc.2019.58.01.20.
170. Vergne-Salle P, Pouplin S, Trouvin AP, Bera-Louville A, Soubrier M, Richez C, et al. The burden of pain in rheumatoid arthritis: Impact of disease activity and psychological factors. Eur J Pain. 2020 Nov;24(10):1979-89. DOI: 10.1002/ejp.1651.
171. Carpenter L, Barnett R, Mahendran P, Nikiphorou E, Gwinnutt J, Verstappen S, et al. Secular changes in functional disability, pain, fatigue and mental well-being in early rheumatoid arthritis. A longitudinal meta-analysis. Semin Arthritis Rheum. 2020 Apr;50(2):209-19. DOI: 10.1016/j.semarthrit.2019.08.006.
172. Walsh DA, McWilliams DF. Pain in rheumatoid arthritis. Curr Pain Headache Rep. 2012 Dec;16(6):509-17. DOI: 10.1007/s11916-012-0303-x.
173. Forslind K, Hafstrom I, Ahlmen M, Svensson B, Group BS. Sex: a major predictor of remission in early rheumatoid arthritis? Ann Rheum Dis. 2007 Jan;66(1):46-52. DOI: 10.1136/ard.2006.056937.
174. Welsing PM, Fransen J, van Riel PL. Is the disease course of rheumatoid arthritis becoming milder? Time trends since 1985 in an inception cohort of early rheumatoid arthritis. Arthritis Rheum. 2005 Sep;52(9):2616-24. DOI: 10.1002/art.21259.
175. Conner TS, Tennen H, Fleeson W, Barrett LF. Experience Sampling Methods: A Modern Idiographic Approach to Personality Research. Soc Personal Psychol Compass. 2009 May 1;3(3):292-313. DOI: 10.1111/j.1751-9004.2009.00170.x.
176. Conaghan PG, Kloppenburg M, Schett G, Bijlsma JW, committee Eoah. Osteoarthritis research priorities: a report from a EULAR ad hoc expert committee. Ann Rheum Dis. 2014 Aug;73(8):1442-5. DOI: 10.1136/annrheumdis-2013-204660.
177. Wieczorek M, Rotonda C, Guillemin F, Rat AC. What Have We Learned From Trajectory Analysis of Clinical Outcomes in Knee and Hip Osteoarthritis Before Surgery? Arthritis Care Res (Hoboken). 2020 Dec;72(12):1693-702. DOI: 10.1002/acr.24069.
178. Allen KD. The value of measuring variability in osteoarthritis pain. J Rheumatol. 2007 Nov;34(11):2132-3.
179. 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. DOI: 10.1002/art.21407.
180. Arksey H OML. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8:19-32.

181. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci. 2010 Sep 20;5:69. DOI: 10.1186/1748-5908-5-69.
182. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018 Oct 2;169(7):467-73. DOI: 10.7326/M18-0850.
183. Affleck G, Tennen H, Urrows S, Higgins P. Individual differences in the day-to-day experience of chronic pain: a prospective daily study of rheumatoid arthritis patients. Health Psychol. 1991;10(6):419-26. DOI: 10.1037//0278-6133.10.6.419.
184. Affleck G, Urrows S, Tennen H, Higgins P. Daily coping with pain from rheumatoid arthritis: patterns and correlates. Pain. 1992 Nov;51(2):221-9. DOI: 10.1016/0304-3959(92)90263-B.
185. Bailly F JC, Servy H, Fautrel B, Gossec L. Variability of pain levels is explained by self-reported disease acitvity in rheumatoid arthritis and axial spondyloarthritis: A 3-months prospective study of 165 patients. Annals of Rheumatic Diseases 2018;77:538.
186. Cruise CE, Broderick J, Porter L, Kaell A, Stone AA. Reactive effects of diary self-assessment in chronic pain patients. Pain. 1996 Oct;67(2-3):253-8. DOI: 10.1016/0304-3959(96)03125-9.
187. Drouet J, Gossec L, Jacquemin C, Fautrel B, Foltz V, Gandjbakhch F, et al. Fluctuation of pain is frequent in rheumatoid arthritis and axial spondyloarthritis: A 12 weeks prospective study of 165 patients. Joint Bone Spine. 2022 May;89(3):105306. DOI: 10.1016/j.jbspin.2021.105306.
188. Eberhard A, Bergman S, Mandl T, Olofsson T, Rydholm M, Jacobsson L, et al. Predictors of unacceptable pain with and without low inflammation over 5 years in early rheumatoid arthritis-an inception cohort study. Arthritis Res Ther. 2021 Jun 14;23(1):169. DOI: 10.1186/s13075-021-02550-7.
189. Finan PH HE, Amazeen PG, Butner J, Zautra AJ, Tennen H. Oscillations in daily pain prediction accuracy. Nonlinear dynamics, psychology and life sicences. 2010;14(1):27-46.
190. 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. DOI: 10.1093/rap/rkac021.
191. Harries L MD, Young A, Kiely PDW, Walsh DA. Prediction of future pain by DAS28-P in patients with early rheumatoid arthritis: The ERAN cohort. Rheumatology (UK). 2014;1:i71-i2.
192. Kumaradev S, Roux C, Sellam J, Perrot S, Pham T, Dugravot A, et al. Socio-demographic determinants in the evolution of pain in inflammatory rheumatic diseases: results from ESPOIR and DESIR cohorts. Rheumatology (Oxford). 2022 Apr 11;61(4):1496-509. DOI: 10.1093/rheumatology/keab562.
193. McWilliams DF, Dawson O, Young A, Kiely PDW, Ferguson E, Walsh DA. Discrete Trajectories of Resolving and Persistent Pain in People With Rheumatoid Arthritis Despite Undergoing Treatment for Inflammation: Results From Three UK Cohorts. J Pain. 2019 Jun;20(6):716-27. DOI: 10.1016/j.jpain.2019.01.001.
194. Odegard S, Finset A, Mowinckel P, Kvien TK, Uhlig T. Pain and psychological health status over a 10-year period in patients with recent onset rheumatoid arthritis. Ann Rheum Dis. 2007 Sep;66(9):1195-201. DOI: 10.1136/ard.2006.064287.
195. Roche PA, Klestov AC, Heim HM. Description of stable pain in rheumatoid arthritis: a 6 year study. J Rheumatol. 2003 Aug;30(8):1733-8.
196. Sandberg M SS, Altawil R, Alfredsson L, Klareskog L, Lampa J. The relationship between pain and inflammation in early rheumatoid arthritis: Long-term follow-up

in the Swedish population-based EIRA and SRQ register. Scandinavian Journal of Rheumatology. 2014;127:4-5.
197. 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):813-22. DOI: 10.1016/j.pain.2012.01.001.
198. Schneider S, Junghaenel DU, Ono M, Stone AA. Temporal dynamics of pain: an application of regime-switching models to ecological momentary assessments in patients with rheumatic diseases. Pain. 2018 Jul;159(7):1346-58. DOI: 10.1097/j.pain.0000000000001215.
199. Smith BW, Zautra AJ. The effects of anxiety and depression on weekly pain in women with arthritis. Pain. 2008 Aug 31;138(2):354-61. DOI: 10.1016/j.pain.2008.01.008.
200. van Dartel SA, Repping-Wuts JW, van Hoogmoed D, Bleijenberg G, van Riel PL, Fransen J. Association between fatigue and pain in rheumatoid arthritis: does pain precede fatigue or does fatigue precede pain? Arthritis Care Res (Hoboken). 2013 Jun;65(6):862-9. DOI: 10.1002/acr.21932.
201. Ward MM. Are patient self-report measures of arthritis activity confounded by mood? A longitudinal study of patients with rheumatoid arthritis. J Rheumatol. 1994 Jun;21(6):1046-50.
202. Ward MM, Leigh JP. The relative importance of pain and functional disability to patients with rheumatoid arthritis. J Rheumatol. 1993 Sep;20(9):1494-9.
203. Wolfe F, Michaud K. Assessment of pain in rheumatoid arthritis: minimal clinically significant difference, predictors, and the effect of anti-tumor necrosis factor therapy. J Rheumatol. 2007 Aug;34(8):1674-83.
204. Lotsch J, Alfredsson L, Lampa J. Machine-learning-based knowledge discovery in rheumatoid arthritis-related registry data to identify predictors of persistent pain. Pain. 2020 Jan;161(1):114-26. DOI: 10.1097/j.pain.0000000000001693.
205. Bastick AN, Wesseling J, Damen J, Verkleij SP, Emans PJ, Bindels PJ, et al. 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):e32-9. DOI: 10.3399/bjgp15X688129.
206. Felson D, Niu J, Sack B, Aliabadi P, McCullough C, Nevitt MC. Progression of osteoarthritis as a state of inertia. Ann Rheum Dis. 2013 Jun;72(6):924-9. DOI: 10.1136/annrheumdis-2012-201575.
207. Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SM, Boers M, Cats HA, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis. 2009 Sep;68(9):1413-9. DOI: 10.1136/ard.2008.096164.
208. Losina E, Collins JE. Forecasting the future pain in hip OA: can we rely on pain trajectories? Osteoarthritis Cartilage. 2016 May;24(5):765-7. DOI: 10.1016/j.joca.2016.01.989.
209. Kim TK. Practical statistics in pain research. Korean J Pain. 2017 Oct;30(4):243-9. DOI: 10.3344/kjp.2017.30.4.243.
210. Haefeli M, Elfering A. Pain assessment. Eur Spine J. 2006 Jan;15 Suppl 1(Suppl 1):S17-24. DOI: 10.1007/s00586-005-1044-x.
211. Peacock S, Patel S. Cultural Influences on Pain. Rev Pain. 2008 Mar;1(2):6-9. DOI: 10.1177/204946370800100203.
212. van Vollenhoven RF. Sex differences in rheumatoid arthritis: more than meets the eye. BMC Med. 2009 Mar 30;7:12. DOI: 10.1186/1741-7015-7-12.

213. Wolfe AM, Kellgren JH, Masi AT. The epidemiology of rheumatoid arthritis: a review. II. Incidence and diagnostic criteria. Bull Rheum Dis. 1968 Nov;19(3):524-9.
214. Maranini B, Bortoluzzi A, Silvagni E, Govoni M. Focus on Sex and Gender: What We Need to Know in the Management of Rheumatoid Arthritis. J Pers Med. 2022 Mar 20;12(3). DOI: 10.3390/jpm12030499.
215. Sokka T, Toloza S, Cutolo M, Kautiainen H, Makinen H, Gogus F, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. Arthritis Res Ther. 2009;11(1):R7. DOI: 10.1186/ar2591.
216. Stone AA, Obbarius A, Junghaenel DU, Wen CKF, Schneider S. High-resolution, field approaches for assessing pain: Ecological Momentary Assessment. Pain. 2021 Jan;162(1):4-9. DOI: 10.1097/j.pain.0000000000002049. Cited in: Medline.
217. Martini CH, Yassen A, Krebs-Brown A, Passier P, Stoker M, Olofsen E, et al. A novel approach to identify responder subgroups and predictors of response to low- and high-dose capsaicin patches in postherpetic neuralgia. Eur J Pain. 2013 Nov;17(10):1491-501. DOI: 10.1002/j.1532-2149.2013.00329.x. Cited in: Medline.
218. Treister R, Honigman L, Lawal OD, Lanier RK, Katz NP. A deeper look at pain variability and its relationship with the placebo response: results from a randomized, double-blind, placebo-controlled clinical trial of naproxen in osteoarthritis of the knee. Pain. 2019 Jul;160(7):1522-8. DOI: 10.1097/j.pain.0000000000001538. Cited in: Medline.
219. Mun CJ, Suk HW, Davis MC, Karoly P, Finan P, Tennen H, et al. Investigating intraindividual pain variability: methods, applications, issues, and directions. Pain. 2019 Nov;160(11):2415-29. DOI: 10.1097/j.pain.0000000000001626.
220. Schneider S, Junghaenel DU, Ono M, Broderick JE, Stone AA. III. Detecting Treatment Effects in Clinical Trials With Different Indices of Pain Intensity Derived From Ecological Momentary Assessment. J Pain. 2021 Apr;22(4):386-99. DOI: 10.1016/j.jpain.2020.10.003. Cited in: Medline.
221. Stone AA, Broderick JE, Schneider S, Schwartz JE. Expanding options for developing outcome measures from momentary assessment data. Psychosom Med. 2012 May;74(4):387-97. DOI: 10.1097/PSY.0b013e3182571faa. Cited in: Medline.
222. Dekker J. Harmonization of pain scores - interesting and relevant findings. Osteoarthritis Cartilage. 2023 Jan;31(1):9-10. DOI: 10.1016/j.joca.2022.08.017.
223. Georgopoulos V, Smith S, McWilliams DF, Steultjens MPM, Williams A, Price A, et al. Harmonising knee pain patient-reported outcomes: a systematic literature review and meta-analysis of Patient Acceptable Symptom State (PASS) and individual participant data (IPD). Osteoarthritis Cartilage. 2023 Jan;31(1):83-95. DOI: 10.1016/j.joca.2022.08.011.
224. Abdullah N. Placebo effect in the treatment of rheumatoid arthritis: a systematic review and meta-analysis of randomized controlled trials. Paper presented at: EULAR ARD; 2015.
225. Bechman K, Yates M, Norton S, Cope AP, Galloway JB. Placebo Response in Rheumatoid Arthritis Clinical Trials. J Rheumatol. 2020 Jan;47(1):28-34. DOI: 10.3899/jrheum.190008.
226. Singer JD, Willett JB. Applied Longitudinal Data Analysis: Modelling Change and Event Occurence. Oxford, New York: Oxford University Press; 2003.
227. 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). DOI: 10.3390/jcm9092828.

228. Bellamy N, Sothern RB, Campbell J, Buchanan WW. Rhythmic variations in pain, stiffness, and manual dexterity in hand osteoarthritis. Ann Rheum Dis. 2002 Dec;61(12):1075-80. DOI: 10.1136/ard.61.12.1075.
229. Jamison RN, Brown GK. Validation of hourly pain intensity profiles with chronic pain patients. Pain. 1991 May;45(2):123-8. DOI: 10.1016/0304-3959(91)90176-X.
230. Murphy SL, Smith DM, Clauw DJ, Alexander NB. The impact of momentary pain and fatigue on physical activity in women with osteoarthritis. Arthritis Rheum. 2008 Jun 15;59(6):849-56. DOI: 10.1002/art.23710.
231. Parrish BP, Zautra AJ, Davis MC. The role of positive and negative interpersonal events on daily fatigue in women with fibromyalgia, rheumatoid arthritis, and osteoarthritis. Health Psychol. 2008 Nov;27(6):694-702. DOI: 10.1037/0278-6133.27.6.694.
232. Stone AA, Broderick JE, Porter LS, Kaell AT. The experience of rheumatoid arthritis pain and fatigue: examining momentary reports and correlates over one week. Arthritis Care Res. 1997 Jun;10(3):185-93. DOI: 10.1002/art.1790100306.
233. Axen I, Bodin L, Bergstrom G, Halasz L, Lange F, Lovgren 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 17;12:99. DOI: 10.1186/1471-2474-12-99.
234. Chen C, Hogg-Johnson S, Smith P. The recovery patterns of back pain among workers with compensated occupational back injuries. Occup Environ Med. 2007 Aug;64(8):534-40. DOI: 10.1136/oem.2006.029215.
235. Deyo RA, Bryan M, Comstock BA, Turner JA, Heagerty P, Friedly J, et al. Trajectories of symptoms and function in older adults with low back disorders. Spine (Phila Pa 1976). 2015 Sep 1;40(17):1352-62. DOI: 10.1097/BRS.0000000000000975.
236. Downie AS, Hancock MJ, Rzewuska M, Williams CM, Lin CC, Maher CG. Trajectories of acute low back pain: a latent class growth analysis. Pain. 2016 Jan;157(1):225-34. DOI: 10.1097/j.pain.0000000000000351.
237. Dunn KM, Campbell P, Jordan KP. Long-term trajectories of back pain: cohort study with 7-year follow-up. BMJ Open. 2013 Dec 11;3(12):e003838. DOI: 10.1136/bmjopen-2013-003838.
238. Dunn KM, Jordan K, Croft PR. Characterizing the course of low back pain: a latent class analysis. Am J Epidemiol. 2006 Apr 15;163(8):754-61. DOI: 10.1093/aje/kwj100.
239. Kent P, Kongsted A. Identifying clinical course patterns in SMS data using cluster analysis. Chiropr Man Therap. 2012 Jul 2;20(1):20. DOI: 10.1186/2045-709X-20-20.
240. 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 1;15(5):885-94. DOI: 10.1016/j.spinee.2015.02.012.
241. 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):210-21. DOI: 10.2522/ptj.20120416.
242. Tamcan O, Mannion AF, Eisenring C, Horisberger B, Elfering A, Muller U. The course of chronic and recurrent low back pain in the general population. Pain. 2010 Sep;150(3):451-7. DOI: 10.1016/j.pain.2010.05.019.
243. 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):622-30. DOI: 10.1016/j.joca.2014.03.009.

244. Dai Z, Lu N, Niu J, Felson DT, Zhang Y. Dietary Fiber Intake in Relation to Knee Pain Trajectory. Arthritis Care Res (Hoboken). 2017 Sep;69(9):1331-9. DOI: 10.1002/acr.23158.
245. Halilaj E, Le Y, Hicks JL, Hastie TJ, Delp SL. Modeling and predicting osteoarthritis progression: data from the osteoarthritis initiative. Osteoarthritis Cartilage. 2018 Dec;26(12):1643-50. DOI: 10.1016/j.joca.2018.08.003.
246. 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):2041-50. DOI: 10.1016/j.joca.2014.09.026.
247. Trouvin AP, Marty M, Goupille P, Perrot S. Determinants of daily pain trajectories and relationship with pain acceptability in hip and knee osteoarthritis. A national prospective cohort study on 886 patients. Joint Bone Spine. 2019 Mar;86(2):245-50. DOI: 10.1016/j.jbspin.2018.06.009.
248. Wesseling J, Bastick AN, ten Wolde S, Kloppenburg M, Lafeber FP, Bierma-Zeinstra SM, et al. 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):1470-7. DOI: 10.3899/jrheum.141036.
249. Cloudy with a Chance of Pain on BBC North West Tonight. 2016 [Accessed 5 April 2023]. Available from: https://www.youtube.com/watch?v=YUdtKGr49GY
250. 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;2:105. DOI: 10.1038/s41746-019-0180-3.
251. Reade S, Spencer K, Sergeant JC, Sperrin M, Schultz DM, Ainsworth J, et al. Cloudy with a Chance of Pain: Engagement and Subsequent Attrition of Daily Data Entry in a Smartphone Pilot Study Tracking Weather, Disease Severity, and Physical Activity in Patients With Rheumatoid Arthritis. JMIR Mhealth Uhealth. 2017 Mar 24;5(3):e37. DOI: 10.2196/mhealth.6496.
252. limited u. uMotif. 2012-2023 [31 March 2023]. Available from: umotif.com
253. 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 1;5(11):e168. DOI: 10.2196/mhealth.8162.
254. Schultz DM, Beukenhorst AL, Yimer BB, Cook L, Pisaniello HL, House T, et al. Weather Patterns Associated with Pain in Chronic-Pain Sufferers. Bulletin of the American Meteorological Society. 2020:E555-66. DOI: https://doi.org/10.1175/BAMS-D-19-0265.1.
255. 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-Feb;7(1):e963. DOI: 10.1097/PR9.0000000000000963.
256. Curran PJ, Bauer DJ. The disaggregation of within-person and between-person effects in longitudinal models of change. Annu Rev Psychol. 2011;62:583-619. DOI: 10.1146/annurev.psych.093008.100356.
257. Bates D, Machler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software. 2015;67(1):1-48. DOI: https://doi.org/10.18637/jss.v067.i01.
258. Grinstead CM, Snell JL. Introduction to Probability: American Mathematical Society; 1997.
259. Tolver A. An Introduction to Markov Chains; for *Stochastic Processes* course. Department of Mathematical Sciences, University of Copenhagen2016.

260. Helske S, Helske J. Mixture Hidden Markov Models for Sequence Data: The seqHMM Package in R. Journal of Statistical Software. 2019;88(1-32). DOI: https://doi.org/10.18637/jss.v088.i03.
261. Beukenhorst AL, McBeth J, Sergeant JC, Schultz DM, Dixon WG. The prevalence of pain flares: impact of definition. Paper presented at: EULAR conference 2018; 2018.
262. Maugars Y, Berthelot JM, Le Goff B, Darrieutort-Laffite C. Fibromyalgia and Associated Disorders: From Pain to Chronic Suffering, From Subjective Hypersensitivity to Hypersensitivity Syndrome. Front Med (Lausanne). 2021;8:666914. DOI: 10.3389/fmed.2021.666914.
263. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. Int J Epidemiol. 2005 Feb;34(1):215-20. DOI: 10.1093/ije/dyh299.
264. Nesselroade JR, Stigler SM, Baltes PB. Regression Toward the Mean and the Study of Change. Psychological Bulletin. 1980;88(3):622-37.
265. Axen I, Bodin L. Searching for the optimal measuring frequency in longitudinal studies -- an example utilizing short message service (SMS) to collect repeated measures among patients with low back pain. BMC Med Res Methodol. 2016 Sep 13;16(1):119. DOI: 10.1186/s12874-016-0221-4.
266. Bayman EO, Oleson JJ, Rabbitts JA. AAAPT: Assessment of the Acute Pain Trajectory. Pain Med. 2021 Mar 18;22(3):533-47. DOI: 10.1093/pm/pnaa440.
267. Hedeker DR, Gibbons RD. Longitudinal Data Analysis. Hoboken, NJ: Wiley-Interscience; 2006.
268. Kannampallil T, Galanter WL, Falck S, Gaunt MJ, Gibbons RD, McNutt R, et al. Characterizing the pain score trajectories of hospitalized adult medical and surgical patients: a retrospective cohort study. Pain. 2016 Dec;157(12):2739-46. DOI: 10.1097/j.pain.0000000000000693.
269. Long JD. Longitudinal Data Analysis for the Behavioral Sciences Using R. Thousand Oaks, CA: SAGE; 2012.
270. Panik MJ. Growth Curve Modeling: Theory and Applications, 1st Edition. Hoboken, NJ: John Wiley & Sons; 2014.
271. Proudman SM, James MJ, Spargo LD, Metcalf RG, Sullivan TR, Rischmueller M, et al. Fish oil in recent onset rheumatoid arthritis: a randomised, double-blind controlled trial within algorithm-based drug use. Ann Rheum Dis. 2015 Jan;74(1):89-95. DOI: 10.1136/annrheumdis-2013-204145.
272. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum. 1988 Mar;31(3):315-24. DOI: 10.1002/art.1780310302.
273. Buchbinder R, March L, Lassere M, Briggs AM, Portek I, Reid C, et al. Effect of treatment with biological agents for arthritis in Australia: the Australian Rheumatology Association Database. Intern Med J. 2007 Sep;37(9):591-600. DOI: 10.1111/j.1445-5994.2007.01431.x.
274. Williams MP, Buchbinder R, March L, Lassere M. The Australian Rheumatology Association Database (ARAD). Semin Arthritis Rheum. 2011 Feb;40(4):e2-3. DOI: 10.1016/j.semarthrit.2010.10.001.
275. Care AGDoHaA. The Pharmaceutical Benefits Scheme (PBS): About the PBS. [updated 1 January 2023]. Available from: https://www.pbs.gov.au/info/about-the-pbs
276. Australia AGS. Rheumatoid arthritis - initial authority application form (PB109). [updated 23 September 2022]. Available from: https://www.servicesaustralia.gov.au/pb109

277. Macfarlane GJ, Barnish MS, Jones GT. Persons with chronic widespread pain experience excess mortality: longitudinal results from UK Biobank and meta-analysis. Ann Rheum Dis. 2017 Nov;76(11):1815-22. DOI: 10.1136/annrheumdis-2017-211476. Cited in: Medline.
278. Walsh DA, McWilliams DF. Mechanisms, impact and management of pain in rheumatoid arthritis. Nat Rev Rheumatol. 2014 Oct;10(10):581-92. DOI: 10.1038/nrrheum.2014.64.
279. Pisaniello HL, Lester S, Russell O, Black R, Tieu J, Richards B, et al. Trajectories of self-reported pain-related health outcomes and longitudinal effects on medication use in rheumatoid arthritis: a prospective cohort analysis using the Australian Rheumatology Association Database (ARAD). RMD Open. 2023 Jul;9(3). DOI: 10.1136/rmdopen-2022-002962.
280. Dadoun S, Zeboulon-Ktorza N, Combescure C, Elhai M, Rozenberg S, Gossec L, et al. Mortality in rheumatoid arthritis over the last fifty years: systematic review and meta-analysis. Joint Bone Spine. 2013 Jan;80(1):29-33. DOI: 10.1016/j.jbspin.2012.02.005.
281. Ferreira RJO, Carvalho PD, Ndosi M, Duarte C, Chopra A, Murphy E, et al. Impact of Patient's Global Assessment on Achieving Remission in Patients With Rheumatoid Arthritis: A Multinational Study Using the METEOR Database. Arthritis Care Res (Hoboken). 2019 Oct;71(10):1317-25. DOI: 10.1002/acr.23866.
282. Studenic P, Smolen JS, Aletaha D. Near misses of ACR/EULAR criteria for remission: effects of patient global assessment in Boolean and index-based definitions. Ann Rheum Dis. 2012 Oct;71(10):1702-5. DOI: 10.1136/annrheumdis-2012-201519.
283. 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. DOI: 10.1016/j.rdc.2019.01.004.
284. Schwartz CE, Sprangers MA. Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. Soc Sci Med. 1999 Jun;48(11):1531-48. DOI: 10.1016/s0277-9536(99)00047-7.
285. Thomas MJ. Expanding and explaining symptoms in knee osteoarthritis trajectories: fluctuations, flares, and future directions. Osteoarthritis Cartilage. 2023 Apr 4. DOI: 10.1016/j.joca.2023.03.014.
286. Black RJ, Richards B, Lester S, Buchbinder R, Barrett C, Lassere M, et al. Factors associated with commencing and ceasing opioid therapy in patients with rheumatoid arthritis. Semin Arthritis Rheum. 2019 Dec;49(3):351-7. DOI: 10.1016/j.semarthrit.2019.06.003.
287. An Australian Living Guideline for the Pharmacological Management of Inflammatory Arthritis: Opioids for Pain Management. [updated 8 November 2022]. Available from: https://app.magicapp.org/#/guideline/LqRV3n
288. Whittle SL, Richards BL, Husni E, Buchbinder R. Opioid therapy for treating rheumatoid arthritis pain. Cochrane Database Syst Rev. 2011 Nov 9(11):CD003113. DOI: 10.1002/14651858.CD003113.pub3.
289. Lee YC, Kremer J, Guan H, Greenberg J, Solomon DH. Chronic Opioid Use in Rheumatoid Arthritis: Prevalence and Predictors. Arthritis Rheumatol. 2019 May;71(5):670-7. DOI: 10.1002/art.40789.
290. Finckh A, Gilbert B, Hodkinson B, Bae SC, Thomas R, Deane KD, et al. Global epidemiology of rheumatoid arthritis. Nat Rev Rheumatol. 2022 Oct;18(10):591-602. DOI: 10.1038/s41584-022-00827-y.
291. Black RJ, Lester S, Tieu J, Sinnathurai P, Barrett C, Buchbinder R, et al. Mortality Estimates and Excess Mortality in Rheumatoid Arthritis. Rheumatology (Oxford). 2023 Mar 15. DOI: 10.1093/rheumatology/kead106.

292. Duarte C, Santos E, da Silva JAP, Kristianslund EK, Kvien TK, Dougados M, et al. The Patient Experienced Symptom State (PESS): a patient-reported global outcome measure that may better reflect disease remission status. Rheumatology (Oxford). 2020 Nov 1;59(11):3458-67. DOI: 10.1093/rheumatology/keaa149.
293. Rifbjerg-Madsen S, Christensen AW, Boesen M, Christensen R, Danneskiold-Samsoe B, Bliddal H, et al. The course of pain hypersensitivity according to painDETECT in patients with rheumatoid arthritis initiating treatment: results from the prospective FRAME-cohort study. Arthritis Res Ther. 2018 May 30;20(1):105. DOI: 10.1186/s13075-018-1581-4.
294. Tanguay-Sabourin C, Fillingim M, Guglietti GV, Zare A, Parisien M, Norman J, et al. A prognostic risk score for development and spread of chronic pain. Nat Med. 2023 Jul;29(7):1821-31. DOI: 10.1038/s41591-023-02430-4.
295. Duarte C, Ferreira RJO, Santos EJF, da Silva JAP. Treating-to-target in rheumatology: Theory and practice. Best Pract Res Clin Rheumatol. 2022 Mar;36(1):101735. DOI: 10.1016/j.berh.2021.101735.
296. Duarte C, Santos EJF, Ferreira RJO, Kvien TK, Dougados M, de Wit M, et al. Validity and reliability of the EULAR instrument RAID.7 as a tool to assess individual domains of impact of disease in rheumatoid arthritis: a cross-sectional study of 671 patients. RMD Open. 2021 Feb;7(1). DOI: 10.1136/rmdopen-2020-001539.
297. McWilliams DF, Walsh DA. Inflammatory and Noninflammatory Disease Activity in Rheumatoid Arthritis: The Effect of Pain on Personalized Medicine. J Rheumatol. 2023 Jun;50(6):721-3. DOI: 10.3899/jrheum.230158.
298. Wohlfahrt A, Muhammad LN, Song J, Dunlop DD, Neogi T, Bingham CO, 3rd, et al. Pain Mechanisms Associated With Disease Activity in Patients With Rheumatoid Arthritis Treated With Disease-Modifying Antirheumatic Drugs: A Regression Tree Analysis. J Rheumatol. 2023 Jun;50(6):741-7. DOI: 10.3899/jrheum.220500.
299. Wohlfahrt A, Muhammad LN, Song J, Dunlop DD, Bingham CO, 3rd, Bolster MB, et al. Pain Mechanisms Associated With Disease Activity in Patients With Rheumatoid Arthritis Treated With Disease-Modifying Antirheumatic Drugs: A Regression Tree Analysis. J Rheumatol. 2023 Jan 15. DOI: 10.3899/jrheum.220500.
300. McWilliams DF, Walsh DA. Inflammatory and Noninflammatory Disease Activity in Rheumatoid Arthritis: The Effect of Pain on Personalized Medicine. J Rheumatol. 2023 Apr 1. DOI: 10.3899/jrheum.230158.
301. Gwinnutt JM, Wieczorek M, Balanescu A, Bischoff-Ferrari HA, Boonen A, Cavalli G, et al. 2021 EULAR recommendations regarding lifestyle behaviours and work participation to prevent progression of rheumatic and musculoskeletal diseases. Ann Rheum Dis. 2023 Jan;82(1):48-56. DOI: 10.1136/annrheumdis-2021-222020.
302. Rausch Osthoff AK, Niedermann K, Braun J, Adams J, Brodin N, Dagfinrud H, et al. 2018 EULAR recommendations for physical activity in people with inflammatory arthritis and osteoarthritis. Ann Rheum Dis. 2018 Sep;77(9):1251-60. DOI: 10.1136/annrheumdis-2018-213585.