

University of Nottingham
UK | CHINA | MALAYSIA

FATIGUE IN RHEUMATOID ARTHRITIS

Ifesemen Onosi Sylvia,
MBBS, MPH

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

July 2024

DECLARATIONS

I hereby declare that this thesis is the result of original research. It has been conducted substantially by me with the assistance received outlined below. This work has not been accepted for any degree, diploma, or other qualification. All authors and works to which reference has been made are fully acknowledged.

Study design, ethical application, data collection and general administration for the early rheumatoid arthritis network (ERAN) inception cohort study data were collected by the ERAN study team, including Dr Adam Young, Professor Patrick Kiely, Professor David Walsh and ERAN data were provided by Dr Samuel Norton and Dr Daniel F MacWilliams. The data dictionary and scientific protocols utilised in ERAN data analyses were designed by me, under the supervision of Professor David A Walsh and Dr Daniel F. McWilliams, with advice from Dr Samuel Norton.

Study design, ethical application, data collection, and general administration for the knee pain and related health in the community (KPIC) study were conducted by the KPIC study team, including Dr Gwen Fernendes, Professor Ana Valdes, Professor Michael Doherty, Professor Weiya Zhang, Dr Michelle Hall and Professor David A Walsh. All scientific protocols utilising KPIC study data in this thesis were designed and conducted by me, under the supervision of Professor David A Walsh and Dr Daniel F McWilliams, with some advice from the KPIC study team.

The study design, design of study materials, including questionnaires, invitation letters, patient and public involvement (PPI) activities and general administration of the CAP-RA study, was done by me under the supervision of Professor David A Walsh and Dr Daniel F McWilliams. In addition, Dr Kehinde Akin-Akinyosoye designed the original CAP-knee questionnaire, and the quantitative sensory testing (QST) protocol was developed primarily by Dr Vasileios Georgopoulos.

Furthermore, members of the CAP-RA study team, including Professor David A Walsh, Dr Daniel F McWilliams, Professor Eamonn Furgerson, Professor Richard Wakefield, Mrs Dorothy Platts, Mrs Susan Ledbury, and Mrs Deborah Wilson contributed to the CAP-RA study design. Dr Julie Jones-Diette and Mrs Debbie Wilson contributed to the CAP-RA PPI event.

All data presented in this thesis were analysed by me, and overall supervision was undertaken by Professor David A Walsh and Dr Daniel F McWilliams.

LIST OF PUBLICATIONS

Published papers related to this thesis

Ifesemen OS, McWilliams DF, Ferguson E, et al. Central Aspects of Pain in Rheumatoid Arthritis (CAP-RA): protocol for a prospective observational study. BMC Rheumatology. 2021 Jun;5(1):23. DOI: 10.1186/s41927-021-00187-2. PMID: 34162435; PMCID: PMC8223274.

Onosi Sylvia Ifesemen, Daniel Frederick McWilliams, Sam Norton, Patrick D W Kiely, Adam Young, David Andrew Walsh, Fatigue in early rheumatoid arthritis: data from the Early Rheumatoid Arthritis Network, Rheumatology, Volume 61, Issue 9, September 2022, Pages 3737–3745, https://doi.org/10.1093/rheumatology/keab947

Conference/other presentations

Fatigue In People With Knee Pain compared to Controls: A Population Based Study. Ifesemen, O.S. et al.Osteoarthritis and Cartilage, Volume 30, S250 - S251. ***Poster presentation at the OARSI World Congress on Osteoarthritis, Berlin, April 2022 (presented on my behalf by Prof. David A Walsh)***

Fatigue in Early Rheumatoid Arthritis Network. 2022 Annual ERAN-ERAS meeting, oral presentation, virtual meeting, March 2022

Fatigue and pain Rheumatoid Arthritis - Oral presentation at the Versus Arthritis, Pain Centre Internal Scientific Meeting – Versus Arthritis Nottingham Pain Centre Internal Scientific Meeting, Nottingham, June 2021.

Fatigue in the General population - Sue Watson Oral presentation event, Nottingham, March 2021. (3rd prize winner)

A Prevalent And Persistent Symptom in Early Rheumatoid Arthritis. Ifesemen O, McWilliams D, Young A, et al. THU0100 Annals of the Rheumatic Diseases 2020;79:264. ***Poster presentation at the 2020 European League Against Rheumatism (EULAR), Berlin. (e-congress due to the Covid-19 pandemic). June 2020***

Pain And Fatigue in Early Rheumatoid Arthritis - A Cross-Sectional Study. Ifesemen, O., McWilliams, D., Young, A., Kiely, P., & Walsh, D. (2019, September). ***Poster presented at 11th Congress of the European Pain Federation EFIC, Valencia, Spain***

Physical Activity and Fatigue in Rheumatic Diseases, a review of evidence – Oral Presentation, Moving your way to health conference, Nottingham, June 2019 (Best student oral presentation).

Fatigue in Rheumatoid Arthritis: UKRime annual showcase, Nottingham, September 2018 (Best elevator pitch winner).

Publications of work not presented in this thesis.

McWilliams, Daniel & Thankaraj, Divya & Jones-Diette, Julie & Morgan, Rheinallt , **Ifesemen, Onosi,** Shenker, Nicholas & Walsh, David. (2021). The efficacy of systemic glucocorticosteroids for pain in rheumatoid arthritis: a systematic literature review and meta-analysis. Rheumatology 2021.

ABSTRACT

Background

Fatigue is a very troublesome symptom in many musculoskeletal diseases. In rheumatoid arthritis (RA), fatigue is one of the most distressing symptoms of the disease. It is not routinely addressed in clinical practice, mainly due to the belief that fatigue is due to chronic inflammation. Fatigue is a core outcome measure in clinical trials and a key symptom whose absence is required to indicate RA remission from the patients' perspective. However, our knowledge of fatigue remains deficient despite increasing RA fatigue research. A better understanding of fatigue is indicated to facilitate early identification and the development of specific interventions targeted at managing fatigue in people with RA. This study aims to increase our understanding of fatigue in RA by unpicking some of the theorized associations between fatigue and other symptoms in rheumatoid arthritis.

Objectives

1. To determine the prevalence as well as characteristics associated with fatigue in RA
2. To describe the longitudinal course of fatigue and examine heterogeneity in the course of fatigue in RA
3. To describe the prevalence and associations of fatigue in the general population
4. Examine the association between fatigue and central sensitisation in RA

Methods

This thesis presents the result of three observational studies related to fatigue (two in an RA population and the other in the general population) and a primary research study protocol. Data from the early rheumatoid arthritis network (ERAN) (N=1236 at baseline) was used to investigate the prevalence, factors associated with, and course of fatigue in RA. Fatigue was measured with the vitality subscale of the SF36 questionnaire (Vitality), normed to the UK population. Vitality was dichotomised, and age/sex standardised to estimate prevalence at baseline. Correlation and linear regression modelling were used to examine baseline associations. Fatigue progression and heterogeneity were described using hierarchical growth curve and group-based trajectory modelling in participants with three or more Vitality measurements (n=729). Finally, baseline characteristics associated with progression heterogeneity were determined with logistic regression.

The prevalence and associations of fatigue in the general population were described using baseline and one-year follow-up data from the knee pain and related health in the community (KPIC) study, a community-based cohort from the East Midlands, N=9513. Fatigue was measured using the vitality sub-scale of the SF12 questionnaire (SF12VT) and central pain mechanisms by pain detection threshold (PPT) collected from the anterior tibia and sternum in a subgroup of participants. Linear and logistic regression analyses were conducted.

Finally, the study design and development of an observational study to examine the association between central sensitisation, pain and fatigue in RA using a simple questionnaire is presented in this thesis.

Results

The estimated RA fatigue prevalence was 45% (CI: 44% - 50%). Female sex, pain, mental health, and disability were associated with fatigue, explaining about half the variability at baseline (R^2 = 0.49). Fatigue remained stable throughout the study period (approximately three years follow-up), exhibiting no clinically significant change. Two distinct fatigue trajectory groups, named 'fatigue' and 'no-fatigue' groups, were identified in the ERAN cohort. The fatigue group comprised 52.6% of the total population and was associated with female sex, pain, mental health, disability, and comorbidity. The area under the receiver operative characteristic curve (AUROC) was 0.81. Inflammation was not associated with fatigue cross-sectionally or longitudinally.

In the general population, the prevalence of fatigue was 45% (CI:35% - 64%). Fatigue was associated with female sex, pain, depression, anxiety, sleep, and comorbidity at baseline (R^2=0.50). Baseline fatigue, pain, depression, sleep, and female sex were associated with fatigue in year one (R^2=0.51). A statistically significant association between fatigue and central sensitisation measured PPT was not demonstrated.

Finally, specifics of the study design, role of patient public involvement, ethics application and research quality control in a longitudinal observational study were presented. The study was set up to involve

baseline and 12-week assessments of 250 participants with rheumatoid arthritis and aimed to investigate factors associated with fatigue, pain and central sensitisation. Assessment tools included validated self-reported instruments, objective markers of inflammation (laboratory and imaging) and different modalities to assess central sensitisation.

Conclusions

Fatigue still represents an unmet need in rheumatoid arthritis and the general population and remains prevalent and persistent in RA. People with specific traits are more likely to report fatigue and may be identified early in the disease course. This sub-population presents opportunities for individualised care packages and represents a target group for further research.

Fatigue as a patient-reported outcome is a cross-cutting symptom with similar associations in many disease conditions. Therefore, further evaluation outside the context of a specific disease is indicated, as this will encourage active knowledge sharing and facilitate a better understanding of fatigue.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisors, Professor David Walsh and Dr Daniel McWilliams, for their support and guidance throughout my PhD. Your encouragement, timely and constructive feedback, open-door approach, and investment in this project are deeply appreciated. Special thanks go to Dr Sam Norton, who provided invaluable statistical support for this project. Thank you for taking the time to review my preliminary results and sending me code snippets.

My sincere thanks go to the staff and PhD students in academic rheumatology, with names ranging from A – Z on the alphabet; your friendship and support enriched the PhD experience. I would like to particularly acknowledge my PhD roomies, Drs Sameer, Subhasisha and Afroditi. I am so grateful for the opportunity to call you friends.

I would like to acknowledge my fellow 'woo' watchers - Angela, Laura and Tom, and my orthopaedic colleagues. You always had a listening ear and provided something to laugh about during the PhD lows.

My special thanks go to the research team and participants of the early rheumatoid arthritis network (ERAN) and the knee pain in the community (KPIC) studies, as well as funders, Versus Arthritis pain centre and ERAN. This work would not be possible without your dedication and time.

I would like to acknowledge the support and friendship of my friends and neighbours. Thanks for your love, friendship, and support – babysitting,

school runs, meals and most of all, encouragement. Yes, Uzo, we did this together.

To my family, I owe you a debt of gratitude. My parents, brothers and sister, cousins, and in-laws. Thank you. I am truly blessed to have you in my corner. Your prayers, support, and absolute confidence in me kept me going.

Emmanuel and my girls. There are not enough words to describe how proud I am to call you mine. Sophia and Victoria, you have given me the most important degree I would ever have, the privilege to be called mummy. Thanks for making every day memorable. It is my greatest blessing to watch you grow. Emmanuel, my love, with tears in my eyes, I say - thank you. You have been my rock.

Finally, to the One who was, is and is to come, my everything, the reason for my being. Thank you. You made all this possible. To You remain all the praise and worship, now and forever. Amen.

‘‘He giveth power to the faint; and to them that have no might he increaseth strength’

Isaiah 40:29 KJV

TABLE OF CONTENT

DECLARATIONS ii
LIST OF PUBLICATIONS iv
ABSTRACT vi
ACKNOWLEDGEMENTS x
TABLE OF CONTENT xiii
TABLE OF FIGURES xvii
LIST OF TABLES xix
LIST OF EQUATIONS xxiv
LIST OF ABBREVIATIONS xxv
1. Introduction to the thesis 1
1.1 Personal statement 1
1.2 Structure of this thesis 2
2. Chapter 2: Background 4
2.1 Introduction to RA 4
2.1.1 History and definition 4
2.1.2 Epidemiology of RA 5
2.1.3 Aetiology, pathogenesis, and risk factors 7
2.1.4 Diagnosing RA 8
2.1.5 Treatment of RA 13
2.1.6 Measuring inflammatory disease activity in RA 14
2.1.7 The concept of early RA 19
2.1.8 RA flares 20
2.2 Introduction to fatigue 21
2.2.1 Definition of fatigue 21
2.2.2 Epidemiology of fatigue 24
2.2.3 Fatigue in Musculoskeletal Diseases 24
2.2.4 Fatigue in RA 25
2.2.5 Treatment of fatigue 43
2.2.6 Summary of literature review 46
2.3 Study rationale 47
2.4 Aim and objectives. 48
2.4.1 Aim 48
2.4.2 Objectives 49
3. Chapter 3 Introduction to methods 50
3.1 Outline 50
3.2 Study datasets 50
3.2.1 The early rheumatoid arthritis network (ERAN) 50
3.2.2 Knee pain and related health in the community (KPIC) study 55
3.3 Statistical Analyses 64
3.3.1 Direct standardisation 64
3.3.2 Correlation and linear regression 66
3.3.3 Linear regression 67
3.3.4 Logistic Regression 70
3.3.5 Growth Models 74

3.3.6 Missing Data 80
3.4 Other methods 82
3.4.1 Non-response bias in population surveys 82
3.4.2 Patient and public involvement in research 84
3.5 Covid 19 Impact 85
3.6 Summary 86

4. A Cross-sectional study of fatigue in rheumatoid arthritis 87

4.1 Outline 87
4.2 Introduction 87
4.3 Aim and objectives 88
4.3.1 Aim 88
4.3.2 Objectives 88
4.4 Methods 89
4.4.1 Study design/source of data 89
4.4.2 Exposures and outcomes 89
4.4.3 Statistical Analysis 91
4.5 Results 93
4.5.1 Prevalence of fatigue at Baseline 95
4.5.2 Baseline characteristics associated with fatigue. 95
4.6 Discussion 106
4.6.1 Key findings and interpretations 106
4.6.2 Strengths and limitations 109
4.7 Conclusion 110
4.8 Summary 110

5. Longitudinal course of fatigue in RA 112

5.1 Outline 112
5.2 Introduction 112
5.3 Aims and Objectives 113
5.3.1 Aim 113
5.3.2 Objectives 114
5.4 Methods 114
5.4.1 Study Design/Source of data 114
5.4.2 Exposures and outcomes 114
5.4.3 Statistical Analysis 114
5.5 Results 119
5.5.1 Overview of the study population 119
5.5.2 Progression of fatigue over time. 124
5.5.3 Trajectories of Vitality 127
5.5.4 Sensitivity Analysis 134
5.6 Discussion 135
5.6.1 Key findings and interpretations 135
5.6.2 Conclusion 141
5.6.3 Summary 141

6. Fatigue in the general population – a knee pain in the community study 142

6.1 Outline 142
6.2 Introduction 142
6.3 Aims and Objectives 144
6.3.1 Aim 144
6.3.2 Objectives 144

6.3.3 Hypothesis 145
6.4 Methods 145
6.4.1 Study design/source of data 145
6.4.2 Exposures and outcomes 145
6.4.3 Statistical Analysis 146
6.5 Results 148
6.5.1 Overview of the study population 148
6.5.2 Prevalence of fatigue in the general population 153
6.5.3 Risk factors for fatigue at baseline. 153
6.5.4 Baseline factors associated with fatigue in year 1 162
6.5.5 Fatigue and central sensitization 174
6.5.6 Secondary Analysis 186
6.6 Discussion 199
6.6.1 Key findings and interpretations 199
6.6.2 Strengths and Limitations 202
6.7 Conclusions 204

7. Fatigue and pain in rheumatoid arthritis: investigating the contribution of central mechanisms 205

7.1 Outline 205
7.2 Introduction 205
7.2.1 Aims and objectives 208
7.3 Methods 209
7.3.1 Study design and setting 209
7.3.2 Ethical considerations 210
7.3.3 Public and patient involvement and engagement in the CAP-RA study 210
7.3.4 CAP-RA study participants 213
7.3.5 Study data collection 216
7.3.6 Statistical analysis 221
7.3.7 Sample size considerations 223
7.3.8 A record of CAP-RA Quality Control - TJC and SJC quality control 224
7.4 Discussion 228
7.4.1 Overview of study aims. 228
7.4.2 Strengths and limitations 229
7.5 Conclusion 231
7.6 Covid 19 Impact 232

8. General discussion 233

8.1 Outline 233
8.2 Overview of thesis aim and objectives. 233
8.3 Summary of Results 234
8.4 Novel findings and implications 237
8.5 Strengths and limitations 239
8.6 Directions for future research 241

9. Conclusion 245

10. References 246

11. Appendices 278

11.1 Publications from PhD 278
11.2 Literature review protocol 280

11.3 Databases 281
11.3.1 ERAN Data Dictionary 281
11.3.2 KPIC data access request form 286
11.3.3 Jyvaskyla Early Rheumatoid Arthritis Data Request 291
11.3.4 Jyvaskyla ethics committee approval 294
11.4 Cross-sectional study of fatigue in RA Appendices 298
11.4.1 Explaining the mediation effect of comorbidities on the association between age and fatigue (ERAN data) 305
11.5 Longitudinal course of fatigue in RA Appendices 307
11.6 KPIC study appendices 323
11.6.1 Weighted KPIC population analysis (Estimating the effects of non-response bias) 326
11.7 CAPRA study appendices 333
11.7.1 SPIRIT checklist 333
11.7.2 CAP-RA Steering Committee TOR 336
11.7.3 Minutes from PPI Focus Group - Kings Mill Hospital 19 February 2020 339
11.7.4 CAP RA study materials 343

TABLE OF FIGURES

Figure 2.1:28 Joints assessed in RA(Tikly, 2009)..... 16
Figure 2.2: Domains of central fatigue..... 23
Figure 2.3: Conceptual model of fatigue in RA(Hewlett et al., 2010) 34
Figure 3.1:KPIC study recruitment flowchart ,source:(Fernandes et al., 2017)..... 58
Figure 5.1: ERAN study participant flow chart..... 120
Figure 5.2: Spaghetti plot showing the progression of Vitality over time in 50 randomly selected ERAN participants. 122
Figure 5.3: Progression of Vitality over the study period in 2 groups. 130
Figure 5.4: AUROC graph of logistic regression model describing baseline predictors of the fatigue- group membership..... 134
Figure 6.1: Flowchart describing KPIC dataset 149
Figure 6.2:Model calibration plot..... 190
Figure 6.3:Area under the ROC curve..... 191
Figure 6.4:Model calibration Plot..... 196
Figure 6.5:Area under the ROC curve..... 197
Figure 7.1: CAP-RA study regimen 215
Appendix figure 11.1: Scatterplot matrix of covariates in the total ERAN population at baseline..... 300
Appendix Figure 11.2: Scatter plot matrix of covariates in the ERAN population with a disease duration of less than 6 months 301
Appendix Figure 11.3:Graphical presentation of residuals from the total ERAN population. 302
Appendix Figure 11.4:Histogram of residuals from the multivariable model of the total ERAN population..... 302
Appendix Figure 11.5:Graphical examination of residuals from the multiple regression model from the ERAN sub-population 303
Appendix Figure 11.6:Histogram of residuals from ERAN participants with a disease duration of less than 6 months..... 303
Appendix figure 11.6.1: Mediation Model (Assumed) 305

Appendix Figure 11.7:Predicted growth curves from 10 randomly selected study participants 309
Appendix Figure 11.8: Trajectories of Vitality by fatigue groups 310
Figure 11.9: Multiple imputation convergence plots..... 321
Figure 11.10: Multiple imputation convergence plots 2..... 322
Appendix Figure 11.11:Histogram of residuals of the multivariable linear regression model of the characteristics associated with SF12Vitality at baseline. 323
Appendix Figure 11.12: Histogram of residuals of the multivariable linear regression model of the characteristics associated with SF12Vitality in year 1..... 325

LIST OF TABLES

Table 2.1: 2010 ACR RA classification criteria 11
Table 2.2:1987 RA classification criteria..... 12
Table 4.1: Baseline characteristics of study population. 94
Table 4.2: Crude and age and sex standardized prevalence of fatigue in the ERAN study population at baseline 95
Table 4.3:Pairwise correlation among baseline characteristics of ERAN population..... 98
Table 4.4: Characteristics associated with Vitality at baseline in the bivariable linear regression analysis. 99
Table 4.5: Characteristics associated with Vitality at baseline in the multiple linear regression analysis. 100
Table 4.6: Characteristics of study participants with disease duration ≤ 6 months..... 102
Table 4.7:ERAN population with a disease duration greater than six months compared with the population with a disease duration of six months or less. 103
Table 4.8: Bivariate linear regression in people with disease duration of ≤ 6 months..... 105
Table 4.9:Multivariable linear regression model in people with disease duration ≤ 6 months. 106
Table 5.1: Descriptive characteristics of the ERAN population with ≥ 3 Vitality measurements (n=729)..... 121
Table 5.2: Table showing the progression of Vitality over time, using a random intercept only, random intercept/random slope, quadratic, and linear spline with internal nodes at the first & second measurement occasion (corresponding to 6 months & 1 year after baseline respectively) 124
Table 5.3:Unconditional and conditional growth model 126
Table 5.4: Comparing fit statistics of 1, 2 and 3 group models. 127
Table 5.5: Fit statistics of each group in the 2-group model 128
Table 5.6: Trajectory estimates of two fatigue groups. 129
Table 5.7: Baseline Characteristics based on Trajectory groups..... 132

Table 5.8: Baseline predictors of Vitality group membership 133
Table 6.1: Overview of total KPIC population at baseline 150
Table 6.2: Descriptive Characteristics of the knee and no-knee pain KPIC subpopulations at baseline 152
Table 6.3: Correlation matrix between SF12Vitality and covariates at baseline in the total KPIC population 154
Table 6.4: Correlation matrix between SF12Vitality and covariates at baseline in people with knee pain (n=2686) 155
Table 6.5: Correlation matrix between SF12Vitality and covariates at baseline in people with no-knee pain (n=6827) 156
Table 6.6:Bivariate linear regression analysis describing the association between SF12Vitality and covariates in the total KPIC population at baseline. 157
Table 6.7: Bivariate regression models describing the association between SF12Vitality and covariates in the knee and no-knee pain subpopulations at baseline. 158
Table 6.8:Multivariable linear regression showing the characteristics associated with fatigue in the total KPIC population. 159
Table 6.9: Multivariable linear regression model showing the characteristics associated with fatigue in the population with and without knee pain. 161
Table 6.10: Characteristics of the KPIC population in year 1 163
Table 6.11: Comparing baseline characteristics between missing and non-missing KPIC participants in year 1. 165
Table 6.12: Correlation of baseline covariates and year 1 SF12Vitality in the total KPIC population 167
Table 6.13: Correlation describing the relationships among baseline characteristics and SF12Vitality in year 1 in people with knee pain and no-knee pain. 167
Table 6.14:Bivariate associations between SF12Vitality in year one and baseline covariates in the total KPIC population 168
Table 6.15: Bivariate associations between baseline characteristics and SF12Vitality in year 1 in the knee and no-knee pain sub-populations 169

Table 6.16: Baseline characteristics associated with year one Vitality in the total KPIC population. 171
Table 6.17: Multivariable linear regression model showing characteristics associated with SF12Vitality in year one, in knee pain and no-knee pain sub-populations..... 172
Table 6.18: Overview of KPIC subpopulation who had additional assessments..... 175
Table 6.19: Overview of baseline characteristics of early and established knee pain groups..... 177
Table 6.20: Table showing the paired correlation matrix for the association between SF12Vitality, PPT, and other variables in the sub-population with knee pain. 179
Table 6.21: Table showing the paired correlation matrix for the association between Vitality, PPT and other variables in the sub-population without knee pain. 180
Table 6.22: Correlation between vitality and covariates in the KPIC population with early knee pain 181
Table 6.23: Correlation between SF12Vitality and covariates in the KPIC population with established knee pain..... 182
Table 6.24: Univariate and multivariable regression examining the association between PPT and SF12Vitality at baseline..... 184
Table 6.25: Association between PPT and Vitality in the no-knee pain sub-population. 185
Table 6.26: Associations between PPT and Vitality in the early and established knee pain subpopulations. 185
Table 6.27: Prevalence of fatigue in the total population and in people with knee pain..... 186
Table 6.28:Overview of people with or without fatigue in the total population..... 187
Table 6.29: Bivariate associations between fatigue and covariates at baseline in the total KPIC population..... 188
Table 6.30: Factors associated with fatigue at baseline in the total KPIC population. 189

Table 6.31: Dominance analysis showing factors associated with fatigue at baseline in ranked order. 190
Table 6.32: Factors associated with fatigue (y/n) at baseline in the knee and no-knee pain KPIC subpopulations. 192
Table 6.33: Table of dominance ranking (knee pain)..... 193
Table 6.34: Table of dominance ranking (no-knee pain) 193
Table 6.35:Bivariate logistic describing the association between baseline covariates and fatigue in year 1 in the total KPIC population. 194
Table 6.36:Baseline covariates associated with fatigue in year 1 in the multivariable logistic regression model in the total KPIC population.. 195
Table 6.37:Dominance analysis showing baseline factors associated with fatigue in ranked order. 196
Table 6.38: Association between fatigue and PPT in KPIC subpopulation who had further assessments. 198
Table 7.1: Descriptive statistics of TJC and SJC for each rater..... 225
Table 7.2: Results of the Two-way random effect interclass correlation coefficient (ICC)..... 226
Appendix Table 11.1:European Standard Population (ESP) 298
Appendix Table 11.2: Test for multicollinearity Total population..... 304
Appendix Table 11.3: Test for multicollinearity population less than 6 months duration..... 304
Appendix Table 11.4: GROLTS Checklist for LGMM..... 307
Appendix Table 11.5:descriptives characteristics of the ERAN population at each measurement occasion 308
Appendix Table 11.6: Fatigue progression over study duration using non-imputed data..... 311
Appendix Table 11.7: Unconditional and conditional growth model unimputed data..... 312
Appendix Table 11.8: Unconditional and fully conditional Growth models, ERAN participants with five complete Vitality measurements and in people with disease duration of six months or less. 313

Appendix Table 11.9: Random effects of models in people with a disease duration of less than six months and participants with five complete measurements. 315
Appendix Table 11.10: Fit statistics of GBTM models for people with disease duration of less than six months and five compete measurements..... 316
Appendix Table 11.11:Fatigue Group membership in people with disease duration of less than 6 months, n=631 317
Appendix Table 11.12: Fatigue Group membership in people with complete measurements(n=208)..... 318
Appendix Table 11.13: Number of groups selected through alternative model specification..... 319
Table 11.14:Comparing East midlands population and the KPIC sample 327
Appendix table 11.15: Comparing Weighted population with KPIC population..... 328
Table 11.16: Bivariate analysis SF12Vitality, KPIC population vs weighted population..... 329
Appendix table 11.17: Multivariable analysis SF12 Vitality, KPIC population vs weighted population..... 330
Appendix table 11.18 Bivariate analysis Fatigue, KPIC population vs weighted population..... 331
Appendix Table 11.19Multivariable analysis Fatigue, KPIC population vs weighted population..... 332

LIST OF EQUATIONS

Equation 2.1: DAS28- SR Formula 18
Equation 3.1: Composition of RDCI 54
Equation 3.2: Direct standardisation formula 65
Equation 3.3: Pearson correlation coefficient formula 66
Equation 3.4: Simple linear regression 67
Equation 3.5: Multiple linear regression equation 69
Equation 3.6: Logistic regression 71
Equation 3.7: Random intercept model 76
Equation 3.8: Random intercept random slope model 77
Equation 7.1: CPM formula 220

LIST OF ABBREVIATIONS

ACPA	Anti-citrullinated protein antibodies
ACR	American college of rheumatology
AIC	Akaike information criterion
APP	Average posterior probability
AUROC	Area under the receiver operative characteristic curve
BIC	Bayesian information criteria
BMI	Body mass index
BRAF	Bristol rheumatoid arthritis Fatigue Scale
CAP-RA	Central aspects of pain in rheumatoid arthritis
CCP	Cyclic citrullinated peptide antibodies
CDAI	Clinical disease activity index
CFA	Confirmatory factor analysis
CFQ	Cognitive failures questionnaire
CFS	Chronic fatigue syndrome
CHF	Congestive heart failure
CI	Confidence interval
CITL	Calibration in the large
COPD	Chronic obstructive pulmonary disease
COSMIN	Consensus-based standards for the selection of health measurement instruments
CPM	Conditioned pain modulation
CRP	C reactive protein
CS	Central sensitisation
CVA	Cerebral vascular accident

DALY	Disability adjusted life year
DAS28ESR	Disease activity score in 28 joints ESR
DI	Disability index
DMARD	Disease modifying anti rheumatoid therapy
DVT	Deep vein thrombosis
ERAN	Early rheumatoid arthritis network
ESP	European standard population
ESR	Erythrocyte sedimentation rate
EULAR	European league against rheumatism
FACIT	Functional assessment of chronic illness therapy fatigue scale
FIML	Full information maximum likelihood
FSS	Fatigue severity scale
GBTM	Group based trajectory modelling
GCM	Growth curve modelling
GMM	Growth mixture modelling
GP	General practitioner
GPPAQ	GP physical activity questionnaire
GRIPPS	Guidance for reporting involvement of patients and the public studies
GROLTS	Guidelines for reporting on latent trajectory studies
HADS	Hospital anxiety and depression scale
HAQ	Health assessment questionnaire
HB	Haemoglobin
HPD	Hepatobiliary disease

HRA	Health research authority
HSE	Health survey for England
HYP	Hypertension
IASP	International association for the study of pain
IBS	Irritable bowel syndrome
ICC	Intraclass correlation coefficient
ICD	International classification of disease
IHD	Ischaemic heart disease
IPAQ	International physical activity questionnaire
IPF	Idiopathic pulmonary fibrosis
IQR	Interquartile range
IRAS	Integrated research application system
JAK	Janus kinase inhibitors
KP	Knee pain
KPIC	Knee pain in the community
LGCA	Latent growth curve analysis
LGMM	Latent growth mixture model
LOCF	Last observation carried forward
MAR	Missing at random
MCAR	Missing completely at random
MCID	Minimum clinically important difference
MCP	Metacarpophalangeal joints
MDQ	Multidimensional questionnaires
MFI	Multi-dimensional fatigue inventory
MH	Mental health

MHAQ	Modified health assessment questionnaire
MI	Multiple imputation
MICE	Multiple imputation using chained equations
MID	Minimum important difference
MNAR	Missing not at random
MOS	Medical outcomes study
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MSK	Musculoskeletal
MTP	Metatarsophalangeal
NHS	National health service
NICE	National institute of clinical excellence
NIHR	National institute of health care research
NREC	Nottingham research ethics committee
NRS	Numerical rating scale
NSAID	Non-steroidal anti-inflammatory
OA	Osteoarthritis
OARSI	Osteoarthritis research society international
OECD	Organization for economic cooperation and development
OHD	Organic heart disease
OLS	Ordinary least squares
OMERACT	Outcome measures in rheumatology
OR	Odds ratio
ONS	Office of national statistics
PACKYRS	Pack years

PALS	Patient advisory liaison service
PCS	Pain Catastrophising scale
PGA	Patient’s global assessment of disease activity
PI	Principal investigator
PIP	Proximal interphalangeal joint
PIS	Participant information sheet
PPI	Patient public involvement
PPT	Pressure pain threshold
PRO	Patient-reported outcome
PROM	Patient-reported outcome measure
PROMIS	Patient-reported outcome measurement information system
PVD	Peripheral vascular disease
QST	Quantitative sensory testing
RA	Rheumatoid arthritis
RAF	Rheumatoid factor
RDCI	Rheumatic disease comorbidity index
RCT	Randomised control trial
REC	Research ethics committee
RF	Rheumatoid factor
ROC	Receiver operating characteristic curve
RR	Risk ratio
SCL	Scleritis/episcleritis
SD	Standard deviation
SDAI	Simplified disease activity index
SDC	Smallest detectable change

SE	Standard error
SEM	Structural equation modelling
SF12VT	SF12 vitality
SF36BP	SF36 bodily pain
SF36MH	SF36 mental health
SF36VT	SF36 vitality
SFH	Sherwood forest hospitals NHS foundation trust
SJC	Swollen joint count
SJS	Sjogren’s syndrome
SLE	Systemic lupus erythematosus
SMD	Standardized mean difference
SMS	Short message service
SYS	Systemic vasculitis
T2T	Treat to target
TJC	Total joint count
TMD	Temporomandibular joint disorder
TNF	Tumour necrosis factor
UK	United Kingdom
VAS	Visual analogue scale
VIF	Variance inflation factor
WPI	Widespread pain index
YLD	Years lived with disability

1. Introduction to the thesis

1.1 Personal statement

At the start of this studentship, I knew very little about the impact of fatigue on many individuals. As a busy mother of two young children working full-time, fatigue played a prominent role on most typical days of my life. During these times, however, I always knew that all I needed to recover was rest. Unfortunately, rest produces no respite for many people who suffer from fatigue in many disease conditions. Through this PhD journey, I have come to appreciate the distress caused by fatigue indirectly by examining data and directly through direct contact with patients and from patient public involvement (PPI) events.

Fatigue is a complex and highly subjective construct related to other multidimensional and complex structures, posing a research dilemma. In the last decade, there has been a proliferation of fatigue-related research; and although this has broadened our understanding of fatigue, the constraints of our knowledge of the condition translate to limited value to the patient. In this PhD, I have drawn from my knowledge and experiences in clinical medicine, epidemiology, medical statistics, and public health to examine fatigue in rheumatoid arthritis. I hope to increase our knowledge of fatigue as it applies to people with RA in the first instance and other disease conditions. Hopefully, the findings from this study move us closer to delivering interventions that positively impact patient care, ultimately improving the quality of life of many people living with RA and other disease conditions.

1.2 Structure of this thesis

The opening chapters of this thesis provide an introduction to the topic and an overview of the methods used in the rest of the thesis. Chapter two introduces rheumatoid arthritis (RA) and fatigue in general and fatigue in RA. This chapter ends with a description of the study's rationale and the aim and objectives of this thesis.

Chapter three introduces methods used in the thesis, including study datasets and statistical methodologies. Theories underpinning the statistical techniques and assumptions governing statistical modelling are discussed. This chapter also includes an overview of missing data and explores a missing data handling technique used in this thesis.

Chapters four, five and six present empirical work using secondary data analysis of the early rheumatoid arthritis network (ERAN) and the knee pain and related health in the community (KPIC) study. Chapters four and five examine the cross-sectional associations and the longitudinal course of RA using ERAN data. Chapter six extends the findings of chapters four and five by investigating fatigue in the general population.

Chapter seven presents the protocol and study development of a longitudinal observational study to validate a self-report questionnaire to measure central sensitisation in RA and its association with pain and fatigue in RA. All empirical chapters include discussion and conclusions sections.

Chapter eight provides a general discussion of findings, pooling together the results and discussions from empirical chapters in the broader

context of the aim and objectives of this thesis. The strengths and limitations of the thesis and areas for future research are discussed. Chapter nine includes the overall conclusion of this thesis. Finally, chapters ten and eleven include references and appendices, respectively.

2. Chapter 2: Background

2.1 Introduction to RA

2.1.1 History and definition

Rheumatoid Arthritis (RA) is a chronic inflammatory disease characterised by joint swelling, tenderness, and destruction of synovial joints leading to loss of function and disability (Aletaha et al., 2010).

The origins of RA are a subject of scientific debate and speculation. One school of thought argues that RA is a disease of the new world, transmitted to the old world by explorers and settlers. Others argue that RA existed in the old world, dating back to the pre-Columbian era (Weisman, 2014).

In contemporary medical literature, the description of RA is often attributed to Augustin-Jacob Landre-Beauvais's doctoral dissertation "Goutte Asthénique Primitive", published in 1800, even though Thomas Sydenham's earlier publication in 1676 clearly described the signs and symptoms of RA. Landre-Beauvais distinguished between what he referred to as 'primary asthenic gout' and general gout, a term used to describe joint disease at the time (Tsoucalas and Sgantzos, 2017).

The term rheumatoid arthritis was introduced by Sir Alfred Barrog in the late 1800s; officially adopted by the British Ministry of Health in 1922, and by the American Rheumatism Association (now the American College of Rheumatology) in 1941.

RA is a syndromic disease consisting of intra and extra-articular features, presenting in a spectrum of severity. Extra-articular features include neuropathy, vasculitis, pleuritis, eye problems (scleritis, episcleritis, ulcerative keratitis), glomerulonephritis, Felty's syndrome, amyloidosis and rheumatoid nodules(Turesson and Jacobsson, 2004). It is responsible for a significant reduction in the quality of life of sufferers. Despite the therapeutic revolution in the treatment of RA in the last two decades, the disease runs a life-long waxing and waning course, and there is no known cure(Weisman, 2011, Weisman, 2014).

2.1.2 Epidemiology of RA

2.1.2.1 Incidence and Prevalence

Rheumatoid Arthritis is one of the most prevalent chronic inflammatory diseases. RA exhibits varying incidence rates across populations. The incidence rates of RA in low and middle-income countries (LMIC) in sub-Saharan Africa are unknown(Rudan et al., 2015). However, in developed countries, on average, the annual incidence rate is highest in North America at 38 (CI:31- 45) cases per 100,000 population and lowest in Southern Europe at 16.5 (CI:9-24) cases per 100,000 population. The reported RA incidence in Northern Europe is 29 (CI:24-36) cases per 100,000. There is some evidence of a decline in RA incidence rates, but this is not consistent worldwide (Minichiello et al., 2016, Abhishek et al., 2017).

A systematic review of 17 studies from several areas of the world revealed significant differences in prevalence rates across the various

study populations, ranging from 3.3 per 1000 in southern Europe to 10.7 per 1000 in North America (Alamanos et al., 2006). Recent studies point to a stable or decreased prevalence in many European countries in the past decade(Abhishek et al., 2017) and in the USA, a reduction of about 1.1 % (CI, 2.0 – 0.2) prevalence in the last decade.

The cumulative incidence of RA is almost twice as high in females than in males. In women, the peak incidence is between the ages of 45 to 74 years. RA is uncommon in males under 45 years; the highest incidence rate in males is over 65 years (Humphreys et al., 2012, Crowson et al., 2011, Novella-Navarro et al., 2021, Cojocaru et al., 2010). Extra-articular disease is, however, commoner in males(Symmons et al., 1994). RA prevalence is also approximately two times higher in females than males, with the highest prevalence of RA in women over 65 years (Humphreys et al., 2012, D'Souza et al., 2010, Cross et al., 2014, Crowson et al., 2011).

RA mortality rates have declined over the years but are higher than that of the general population (Widdifield et al., 2015). RA, however, constitutes a considerable burden to health care systems, individuals and, by extension, the society at large. It has been reported that the costs associated with in-patient care contribute the most significant proportion of direct RA costs, accounting for between 43% and 75% of the annual RA cost (Cooper, 2000). This has likely changed in recent times due to better control of RA with the use of disease-modifying anti-rheumatic therapy leading to the rarity of inpatient management of RA. However, in the US, the annual healthcare costs of the average RA patient are almost three-fold higher than the non-RA patient (Chen et al., 2018).

In the 2010 Global burden of disease study, RA was ranked 74th out of 291 conditions in terms of disease burden measured as disease adjusted life years (DALYs) and 42nd in terms of disability measured by years of life lived with disability (YLD)(Murray et al., 2012). In addition, about 75% of people newly diagnosed with the disease are of working age, and it is a significant cause of work absence due to sickness. It may ultimately lead to early retirement or joblessness(Sokka, 2003, Hansen et al., 2016, Luís et al., 2020). RA is therefore associated with substantial micro and macro-economic impact. RA has been reported to cost the UK economy about £8 billion annually in productivity losses and costs the Italian Government about €3 Billion in direct and indirect costs (Fakhouri et al., 2018).

2.1.3 Aetiology, pathogenesis, and risk factors

Like other systemic autoimmune diseases, RA may result from a breakdown of homeostatic mechanisms that maintain immune tolerance as part of a complex interplay involving environmental factors in genetically susceptible individuals. The exact aetiology of RA is unknown; however, there is evidence of a genetic and epigenetic component to the cause of RA. The genetic susceptibility of the disease has been estimated to range between 12 – 66% in twin studies(van der Woude et al., 2009, Wysocki et al., 2020). In addition, several genetic risk loci associated with susceptibility to autoimmune conditions, including RA, have been identified from cohort studies worldwide.

Several environmental factors have been linked with RA; among these, smoking is the most studied modifiable RA risk factor. A systematic

review identified numerous epidemiological studies from different countries using different study designs that consistently describe smoking as a risk factor for developing RA (Lahiri et al., 2012).

The effect of diet on RA is less consistent; studies have examined the effect of omega-3 fatty acids, antioxidants, red meat, alcohol, and vitamin D on RA. The effects of these factors on RA range from protective through potentiating to no significant effect. Obesity and the use of statins also provided inconsistent results in the literature (Lahiri et al., 2012).

Socioeconomic status, which may be a proxy marker for unmeasured lifestyle variables, has been linked to poorer outcomes in RA(Harrison et al., 2009); however, the risk of developing RA is not consistently associated with socioeconomic status.

2.1.4 Diagnosing RA

Rheumatoid arthritis comprises a broad spectrum of disease presentations ranging from mild, self-limiting disease to florid systemic inflammation, joint destruction, and severe disability. Onset is usually insidious over weeks to months; however, RA has been known to occur rapidly over days(Grassi et al., 1998).

2.1.4.1 Clinical presentation

Initial symptoms could be articular or extra-articular but are usually non-specific, including fatigue, malaise, puffy hands, and diffuse musculoskeletal pain with or without fever. Joint symptoms usually begin in one joint, rapidly progressing to other joints. RA is often characterised by symmetry and multiple joint involvement (symmetrical polyarthritis);

however, an asymmetrical and single-site presentation is not uncommon (Plenge, 2009, Harris, 2001).

RA can affect any joint; however, the metacarpophalangeal (MCP), proximal interphalangeal (PIP) and wrist joints are most commonly affected. Larger joint involvement usually occurs at later stages of the disease, following the earlier mentioned smaller joints. Joints of the thoracic, lumbar, and sacral spine and hip joints are not frequently involved. Clinical examination of these patients with RA may reveal joint swelling, warmth and tenderness(Harris, 2001).

2.1.4.2 Classification criteria

Diagnostic criteria define a group of symptoms, signs and tests used in routine clinical care as guidance for the management of individual patients. Classification criteria, on the other hand, are set definitions with the primary aim of describing a homogenous population for the purpose of clinical research. Classification criteria tend to have high specificity, generally at the expense of lower sensitivity, making them inappropriate for clinical care(Aggarwal et al., 2015). With RA, rheumatologists are often faced with the challenge of distinguishing RA from other diseases with similar signs and symptoms primarily due to the absence of pathognomonic signs and the heterogenous nature of RA. RA diagnosis is therefore a complex and highly cognitive process requiring processing multiple pieces of information beyond simple set criteria. In order to facilitate RA research, however, classification criteria were developed. The first widely accepted RA classification criteria were developed by the American Rheumatism Association (ARA), now called the American

College of Rheumatology (ACR), in 1957. These were reviewed in 1987 by the ACR to improve clinical applicability. According to these criteria, a person is said to have RA if he/she has satisfied at least four of the seven criteria. Components of the 1987 ACR criteria required signs of joint damage, thereby limiting its ability to detect early disease and the possibility of early intervention, which increased the chances of irreversible joint damage and increased disability(Sokolove and Strand, 2010).

Incidence and prevalence rates of RA may be affected by classification criteria. However, although the 2010 classification criteria identify RA earlier at baseline, incidence rates are similar to the cumulative five-year incidence using the 1987 criteria(Humphreys et al., 2012).

The most recent review by the ACR and European League Against Rheumatism (EULAR) in 2010 was done to improve sensitivity in early disease(Kay and Upchurch, 2012). Although the old criteria misclassified people with early disease, the new classification criteria appear to trade off specificity for sensitivity, with the misclassification of 20 - 60% of people with early undifferentiated arthritis(Zeidler, 2012, Zeidler, 2011, van der Linden et al., 2011).

Tables 2-1 and 2-2 below show the 2010 and 1987 ACR Rheumatoid Arthritis classification criteria (Arnett et al., 1988).

Table 2.1: 2010 ACR RA classification criteria

Joint Distribution (0-5)	
One large joint	0
2-10 large joints	1
1-3 small joints (large joints not counted)	2
4 - 10 small joints (large joints not counted)	3
>10 joints (at least one small joint)	5
Serology (0-3)	
Negative RF AND negative ACPA	0
Low positive RF OR low positive ACPA	2
High positive RF OR high positive ACPA	3
Symptom Duration (0-1)	
< 6 weeks	0
≥ 6 weeks	1
Acute Phase Reactants (0-1)	
Normal CRP AND normal ESR	0
Abnormal CRP OR abnormal ESR	1

RF – Rheumatoid Factor, ACPA – **A**nti-citrullinated peptide antibody, CRP – C-Reactive protein, ESR - Erythrocyte Sedimentation Rate

Table 2.2:1987 RA classification criteria

	Criteria	Definition
1.	**Morning stiffness**	Morning stiffness in and around the joints, lasting at least 1 hour before maximal improvement
2.	**Arthritis of three or more joint areas**	At least three joint areas simultaneously have had soft tissue swelling or fluid (not bony overgrowth alone) observed by a physician. The 14 possible areas are right or left PIP, MCP, wrist, elbow, knee, ankle, and MTP joints
3.	**Arthritis of hand joints**	At least 1 area swollen (as defined above) in a wrist, MCP, or PIP joint
4.	**Symmetric arthritis**	Simultaneous involvement of the same joint areas (as defined in two) on both sides of the body (bilateral involvement of PIPS, MCPs, or MTPs is acceptable without absolute symmetry
5.	**Rheumatoid Nodules**	Subcutaneous nodules, over bony prominences, or extensor surfaces, or in juxta-articular regions, observed by a physician
6.	**Serum rheumatoid factor (RF)**	Demonstration of abnormal amounts of serum rheumatoid factor by any method for which the result has been positive in 4% of normal control subjects
7.	**Radiographic changes**	Radiographic changes typical of rheumatoid arthritis on posteroanterior hand and wrist radiographs, which must include erosions or unequivocal bony decalcification localized in or most marked adjacent to the involved joints (osteoarthritis changes alone do not qualify)

PIP – Distal interphalangeal joints; PIPs – Proximal interphalangeal joints; MCPs – Metacarpophalangeal joints; MTP - Metatarsophalangeal joints.

2.1.5 Treatment of RA

The management of RA has evolved over the years, stemming from a better understanding of the pathophysiology of the disease. This has led to a shift in management modalities from steroids and non-steroidal anti-inflammatory drugs (NSAIDs) to conventional disease-modifying anti-rheumatic drugs (DMARDs) and, more recently, the addition of biologic DMARDs, which target specific pro-inflammatory cytokines(Smith et al., 2011).

DMARDs are a heterogeneous group of agents that reduce inflammation and limit joint damage helping to improve function. Examples of DMARDs include Methotrexate, tumour necrotic factor (TNF) inhibitors, and Janus kinase inhibitors (JAK inhibitors).

Traditionally RA was managed using the pyramidal approach based on symptom management, where DMARDs were introduced and doses increased based on disease progression (Sokka et al., 1999). However, several studies showing the benefit of earlier intensive treatment with DMARDs have suggested that there may be a “window of opportunity” where treatment is most effective. This is possibly due to the “stereotyping of the immunopathological processes operating in the joint”, making them similar to that of patients with markedly different disease durations (Raza et al., 2006, Strand et al., 2016b). A systematic review of 22 RA pharmacological guidelines from across the world revealed a common thread on the immediate commencement of DMARD therapy, with biologic DMARDs introduced in disease refractory to methotrexate(Mian et al., 2019).

2.1.6 Measuring inflammatory disease activity in RA

Possibly due to the heterogeneity of RA, there is no gold standard quantitative measure of RA disease activity. This has led to the development of several measures used in clinical care and research to facilitate disease monitoring. These measures often include laboratory, radiographic, clinical and patient-reported assessments (Zatarain and Strand, 2006).

2.1.6.1 Laboratory measures of RA

Laboratory measurement of RA disease activity involves measuring inflammation markers such as erythrocyte sedimentation rate (ESR) and C - reactive protein (CRP). ESR is the rate at which red blood cells (RBCs) fall through plasma when suspended in a vertical tube. It is an indirect measure of fibrinogen, the most abundant acute phase reactant. ESR is affected by the morphology and number of RBCs, age and sex, and systemic and local inflammation. ESR plays a vital role in the diagnosis and management of RA, is included in the 2010 ACR/EULAR RA classification, a component of disease activity score in 28 joints (DAS28) and the simplified disease activity index (SDAI)(Shapiro, 2021).

CRP is an acute phase reactant (secreted as a consequence of tissue injury and inflammation) and is elevated in RA. CRP is associated with both acute and chronic inflammatory states. Elevated CRP is associated with various metabolic stressors including type 2 diabetes, atherosclerosis, sedentary lifestyles, and unhealthy diet. CRP is also affected by age, sex and race though less so than ESR. ESR and CRP

are therefore not specific to RA. In addition, ESR was found to be normal in about 40% of individuals with RA at the first visit(Shapiro, 2021).

Rheumatoid Factor (RF) is another serological marker for RA diagnosis. RF was the only serological marker for RA in the 1987 ACR classification due to its high specificity. A meta-analysis of 50 studies showed a pooled specificity 85% and a sensitivity of 69% in RA, even though RF may be present in the healthy population and other connective tissue diseases such as systemic lupus erythematosus (SLE) and primary Sjogren syndrome. (Nishimura et al., 2007).

Anti-citrullinated Protein (CCP) is the latest addition to RA biomarkers. It has a sensitivity of 67% and a specificity of 95% and can predict the development of RA in patients with early undifferentiated disease and disease severity in established disease. Anti-CCP and RF are included in the new classification of disease (Nielsen et al., 2012, Shen et al., 2015, Nishimura et al., 2007).

2.1.6.2 Clinical measures of RA disease activity

Clinical Markers of RA disease activity often involve the swollen joint count (SJC) and tender joint count (TJC). This is based on the premise that inflamed joints should be swollen and tender. Over the years, the number of joints counted has evolved from 86 peripheral joints to 28 joint count scores in the 1980s. The 28 joint count is widely accepted and involves assessing ten proximal interphalangeal joints (PIP) of the fingers, ten metacarpophalangeal joints (MCP), wrist, elbows, shoulders, and knees. See figure 2.1 below.

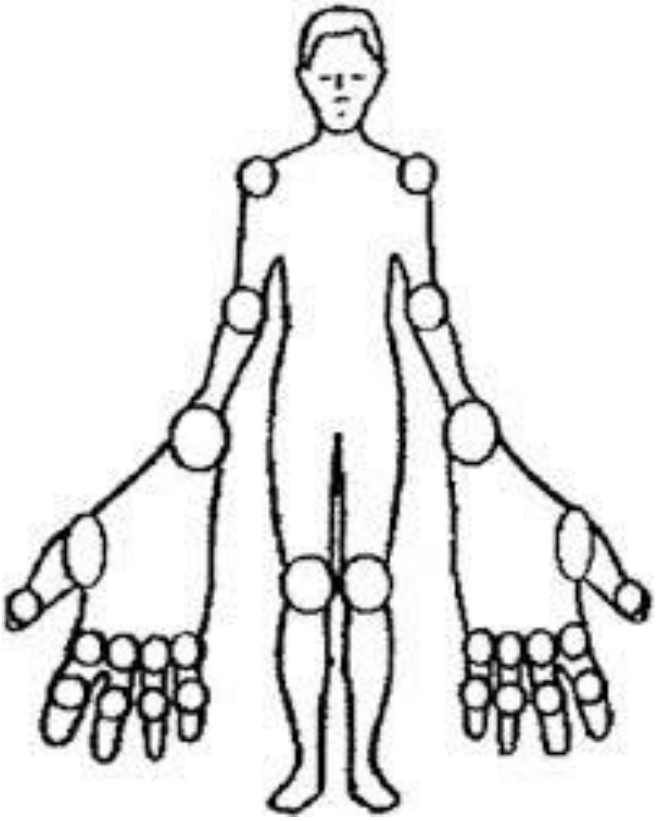


Figure 2.1:28 Joints assessed in RA(Tikly, 2009)

2.1.6.3 Patient-reported measures of RA disease activity

Patient-reported outcomes (PROs) are also used to evaluate RA patients. A PRO is “any report of the status of a patient’s health condition that comes directly from the patient without the interpretation of the patient’s response by a clinical outcome or anyone else”(FDA, 2009). Patient-reported outcome measures (PROMs) are used to measure PROs and are often self-reported questionnaires. Proxy reports or observations from caregivers, health professionals or parents and guardians are strictly not PROMs but clinician or observer-reported outcomes(Powers et al., 2017, Churruca et al., 2021). PROs are pivotal for patient outcomes such as pain, fatigue, quality of life, and emotional function where physiological measurements are limited or unavailable. These attributes hoped to be captured by PROMs are called constructs.

Constructs could be outcomes, as well as risk factors for the development of other conditions(Remes et al., 2021). In this thesis, the word PROMs is used as an umbrella term to describe constructs measured by the patient-reported outcome measure. PROMs could be generic, such as the short-form general health survey (SF36), or disease-specific, such as patient global assessment of disease activity (PGA) and health assessment questionnaire (HAQ). PGA encapsulates the effect of RA on the patient, including inflammation, joint damage, and psychological impact. EULAR and OMERACT endorsed PGA as one of the core sets of outcomes reported in RA clinical trials and is a component of the disease activity score in 28 joints (DAS28). PGA has good psychometric properties but may be affected by heterogeneity in wording/phrasing(Nikiphorou et al., 2016).

These PROMs provide insight into other aspects of the disease that adversely affect the patient's quality of life, such as fatigue, sleep, and the patient's assessment of pain. PROMs are currently used in RA clinical trials to measure the efficacy of interventions and are included in composite scores used to measure RA disease activity (Russell, 2008).

2.1.6.4 Composite measures of RA disease activity

Composite scores were initially developed to facilitate analysis/interpretation of RA disease activity and minimize measurement error in clinical trials and have found their way into routine clinical practice. Several composite scores are available to monitor RA disease activity. ACR recommends disease activity scores for clinical practice; some of these include the clinical disease activity index (CDAI), disease

activity score in 28-joint counts (DAS28) with ESR or CRP, patient activity scale (PAS), PAS-II, routine assessment of patient index data with three measures, and simplified disease activity(Anderson et al., 2012). DAS28 is the most widely used (Kapral et al., 2007).

DAS28 is made up of four RA measures: Tender Joint Count (TJC), Swollen Joint Count (SJC), ESR or CRP and patients' global assessment of disease activity (PGA). It is calculated using the formula below (Fleischmann et al., 2017).

Equation 2.1: DAS28- SR Formula

$$DAS28 = 0.56 \times (\sqrt{(2\&TJC)}) + 0.28 \times (\sqrt{(2\&SJC)}) + 0.70 \times In(ESR) + 0.014 \times GH$$

Where, TJC and SJC are 28 swollen and tender joint counts, respectively, ESR is the erythrocyte sedimentation rate, and GH is the global assessment of disease activity.

SJC and ESR/CRP are often considered objective components, while TJC and PGA are considered subjective components. DAS28 is clinically helpful in assessing a patient's response to treatment and has been extensively validated to evaluate disease activity in clinical trials. Scores range from 0 to 9.4, with values < 2.6 signifying disease remission and ≤ 3.2 defined as low disease activity. Values of 5.1 and above represent high disease activity(Greenmyer et al., 2020). DAS28 can be affected by other factors outside disease activity, including pain perception, sex, and age. In addition, the subjective components of DAS28, particularly PGA, have been found to be highly correlated with cognitive factors and depression(Leeb et al., 2007).

Furthermore, female sex and subjective DAS28 components predicted a lack of pain improvement with treatment in early RA, and fibromyalgia was found to be associated with high DAS28 scores after adjusting for other clinical factors(Ranzolin et al., 2009, McWilliams et al., 2012). This suggests that the subjective components of DAS28 may reflect non-inflammatory factors.

2.1.7 The concept of early RA

RA is often described in terms of disease duration as early and established. The definition of early RA has evolved over the last three decades. Historically, disease duration of five years or less was considered early RA, moving to two years or less in the early 1990s and 12 months in the early 2000s(Scott, 2007). Disease duration of less than three and six months has also been described as early RA; however, this is not universally accepted by rheumatologists (Heidari, 2011). Varying time intervals between symptom onset and diagnosis make the delineation of early versus established RA less objective(Palm and Purinszky, 2005, Aletaha et al., 2002, Gray and Nuki, 2001). This is also complicated by the presence of undifferentiated arthritis, which may be indistinguishable from RA in the early stages(Dixon and Symmons, 2005). Studies show superior clinical, radiological, and functional outcomes to treatment in early than established RA with intensive treatment(Korpela et al., 2004, Jamal et al., 2009). A large Swedish cohort study showed that people with early RA were more likely to remain in sustained remission than those with established disease(Einarsson et al., 2018).

2.1.7.1 Treat to target

The “treat to target” approach, which represents a proactive rather than reactive approach to RA management, was first recommended in 2010 by an international expert committee. Drawing parallels with other chronic diseases like diabetes and hyperlipidaemia; it was highlighted that using treatment targets based on predefined biomarkers and/or other disease activity measures was associated with better clinical outcomes. The committee recommended that remission, defined as the “absence of signs and symptoms of inflammatory disease activity”, should be the primary target of RA treatment, with the concession of low disease activity being the therapeutic goal, especially in patients with joint damage and long-standing disease. The committee also recommended the regular monitoring of disease activity using validated composite measures as well as patient involvement in the treatment process (Smolen and Aletaha, 2011, Toledano et al., 2018).

2.1.8 RA flares

In RA, the term flare is often used to describe a period of worsening disease activity involving a range of signs and symptoms of varying duration and magnitude. Flares could be short-lived and self-limiting, requiring little or no intervention, and could be debilitating, resulting in changes to the pharmacological management of the patient(Bingham et al., 2011). Over the years, several flare definitions have been described based on different indices, including composite RA measures, patients’

perspectives and physician's assessments, and a combination of these(Bozzalla-Cassione et al., 2022).

The Outcome Measures in Rheumatology (OMERACT) RA Flare Group (RA-FG) was established in 2006 with a mandate to determine a validated outcome measure to identify RA flare(Bingham et al., 2009). Activities of this group led to the following definition of RA flare

"*A flare occurs with any worsening of disease activity that would, if persistent, in most cases lead to initiation or change of therapy; and a flare represents a cluster of symptoms of sufficient duration and intensity to require initiation, change or increase in therapy.*"

The group also identified a set of core outcomes required in the definition of RA flare: pain, function, swollen and tender joints, patients' assessment of disease activity, fatigue, stiffness, self-management and participation (Bykerk et al., 2014).

2.2 Introduction to fatigue

2.2.1 Definition of fatigue

The concept of fatigue is difficult to define because of the complex biological, psychosocial, and behavioural interactions involved. As a result, there is no consensus on a universally acceptable definition of fatigue. Authors have attempted to define fatigue in terms of source as central versus peripheral, physical versus psychological, physiological versus pathological, and chronologically as acute versus chronic fatigue (Shen et al., 2006).

Fatigue is a common daily occurrence in healthy people, usually following strenuous physical activity, a busy workday or after a period of emotional tension. In these individuals, symptoms are relieved by rest. The transient nature of this experience means it is rarely considered a serious problem (Tack, 1990).

Fatigue has been classified as central or peripheral. Peripheral fatigue is described in relation to exercise, while central fatigue describes a failure to initiate and sustain tasks requiring self-motivation. Central fatigue can be further categorised into physical and mental fatigue, and motivational changes. Central fatigue could result from an adaptive physiologic process, as described earlier, or it could be pathologic(Karshikoff et al., 2017). Figure 2.2 shows several dimensions of central fatigue(Karshikoff et al., 2017).

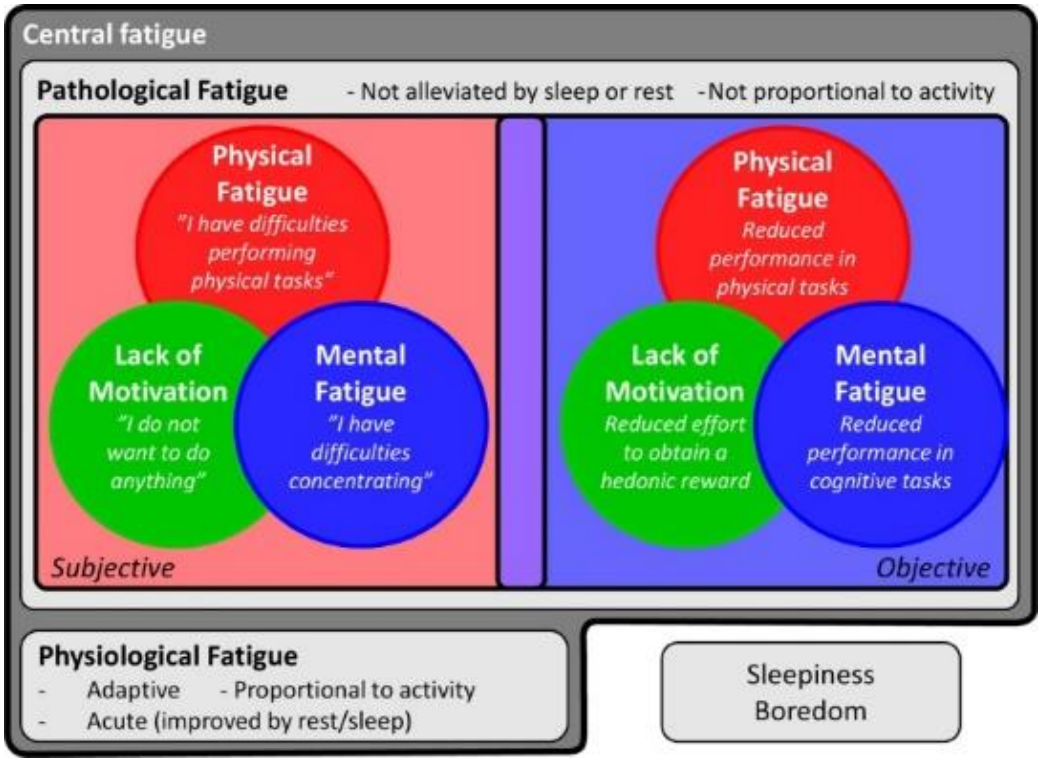


Figure 2.2: Domains of central fatigue

The rest of this thesis describes pathological fatigue, which is not a homeostatic response. Pathologic fatigue (fatigue) describes a state of severe exhaustion, disproportionate to activity and not relieved by rest (Louati and Berenbaum, 2015, Tack, 1990). A more concise definition, widely used in the literature which includes a time element, was proposed by Piper. According to Piper, fatigue can be described as an "*unpleasant, unusual, abnormal or excessive whole-body tiredness, disproportionate to and unrelated to activity and exertion, and present from more than one month*" (Repping-Wuts et al., 2009).

Fatigue is associated with psychiatric, endocrine, inflammatory, haematological, cardiovascular, infectious diseases and several cancers.

People who suffer from fatigue report that the condition is associated with significant distress and disruptions to their way of life (Dupond, 2011, Hewlett et al., 2005).

2.2.2 Epidemiology of fatigue

Fatigue is prevalent in the general population, with prevalence rates reported between 18 – 24% in different populations (Loge et al., 1998a, Pigeon et al., 2003). In primary care, fatigue is cited as the main or secondary reason for clinic visits in 2 - 20% of cases (Cathébras et al., 1992, Fuhrer and Wessely, 1995, Hickie et al., 1999). Women are more likely to report fatigue than men (Pawlikowska et al., 1994).

2.2.3 Fatigue in Musculoskeletal Diseases

Globally, musculoskeletal (MSK) conditions are the most significant contributors to years lived with disability (YLD), accounting for 149 million [CI:108 –199 million] YLDs(Cieza et al., 2020). Fatigue is common in MSK conditions, affecting about 41% of people with osteoarthritis, over 75% of people with Fibromyalgia, and 69% of people with chronic lower back pain(Snekkevik et al., 2014). Fatigue is also prevalent in inflammatory diseases such as RA, psoriatic arthritis, systemic lupus erythematosus (SLE), primary Sjogren Syndrome (pSS)(Meijer et al., 2009, Westhoff et al., 2012, McCormick et al., 2019). In psoriatic arthritis, it is reported to affect one in every two persons with the disease(Overman et al., 2016).

Fatigue is often described in the context of a disease process, for example, post-cancer fatigue, post-stroke fatigue, or post-viral fatigue,

except in the case of chronic fatigue syndrome (myalgic encephalomyelitis), where fatigue is the primary disorder.

2.2.4 Fatigue in RA

The following literature review was generated using a systematic approach. Key search terms, “fatigue”, “tiredness”, “exhaustion”, and rheumatoid arthritis were used according to the protocol described by Brammer et al (Bramer et al., 2018).

Searches were conducted on three databases using the Ovid interface between January 2018 and December 2019. Conference abstracts, non-English articles and non-open access articles inaccessible with the University of Nottingham UK access foundation permission or interlibrary loan service were excluded. Reference lists of included reviewed articles were screened for relevant references according to the screening criteria. Given the broad scope of the information required and the heterogeneity of included studies, a narrative review was conducted (Yuan and Hunt, 2009). Appendix 11.2 contains the detailed framework of the literature search.

2.2.4.1 Introduction

Fatigue exacts a significant influence on the health and quality of life of people with RA. It is associated with worse mental health status, higher levels of interpersonal stress and greater healthcare utilization (Santos et al., 2019). This makes the management of other RA symptoms challenging.

Fatigue is a common occurrence in people with rheumatoid arthritis. However, varying prevalence rates have been reported in research and clinical practice,. This is largely due to differences in definition, measurement indicators, and country of residence (Hifinger et al., 2016). Studies have reported prevalences of between 40% and 80% (Mayoux-Benhamou, 2006, Pinals et al., 1981, Belza, 1995).

2.2.4.2 Fatigue in RA – patients' perspective

RA fatigue often attributed to disease activity, from the patient's perspective. It has been described as different and more intense than any other fatigue episode prior to diagnosis(Hewlett et al., 2010, Crosby, 1991). People with RA fatigue believe that RA related pain and the resulting sleep disturbance are important factors that contribute to the experience of fatigue.

RA related fatigue is described as overwhelming, unpredictable and the worst aspect of the disease permeating through all activities of their lives (Repping-Wuts et al., 2008, Mancuso et al., 2000).

People with fatigue describe psychosocial consequences of fatigue primarily resulting from reduced energy to participate in physical and social activities. They express feelings of fear, hopelessness, frustration, and anger at the impact of fatigue on their daily lives. This is particularly difficult for younger women who occupy multiple roles in society (Repping-Wuts et al., 2008). These individuals describe feelings of inadequacy and uselessness because of the inability to take initiative or meet physical or social expectations. Consequently, from the patient's perspective, fatigue

is associated with social isolation and loneliness (Thomsen et al., 2015, Feldthusen et al., 2013).

A meta-synthesis of qualitative studies of fatigue in RA comprising 212 patients, from different countries identified themes relating to the unpredictability of the symptom, and the inability to communicate the feeling of fatigue in an adequate manner permeating through all included studies –

> *"tired" is not considered an appropriate word, as it is thought not to convey the feeling appropriately*"(Primdahl et al., 2019).

This same thought was expressed during a patient public involvement (PPI) event in which I presented my PhD thesis to a group of people with RA and other inflammatory arthritis.

In a multinational qualitative RA fatigue study based in countries with uniquely different health care systems, a large proportion of respondents could not recall receiving professional advice for fatigue. Support, when provided was generic and not individualised (Repping-Wuts et al., 2008, Dures et al., 2012). Patients have been described as using words and cues relating to mere tiredness that could be relieved by rest or indicating worry or concern about the disease rather than conveying the true nature of the debilitating fatigue they experience (Rijk et al., 2000). Perhaps, patchy professional support could result from suboptimal communication to convey the feeling of fatigue between patients and health professionals or a lack of understanding of the patient's experience, or the "treat the

disease, treat the symptom" assumption made about fatigue by health professionals.

People with RA related fatigue often through trial and error develop coping strategies to manage the symptom. These strategies often require social isolation partly because patients do not feel their symptoms are acknowledged by others. The prioritisation of activities essential to life such as work and immediate needs by these individuals often overtake social relationships leading to a self-perpetuating cycle of isolation and loneliness (Dures et al., 2012, Connolly et al., 2015, Thomsen et al., 2015).

It is interesting to note here that the patient's experience of RA related fatigue and associated impact on quality of life is not unique to RA. Similar descriptors have been used in other chronic conditions such as chronic fatigue syndrome, multiple sclerosis, post viral fatigue, liver-related diseases, such as primary sclerosing cholangitis, primary biliary cirrhosis and non-alcoholic fatty liver disease, and cancer-related fatigue(Jones et al., 2009). It is, however, unclear whether these apparent similarities are superficial or deep, deep in the sense that they reflect the shared experiences of a common pathophysiological process, or superficial, whereby distinct disease processes result in specious similarities in patient experiences and impact. Another related explanation could be a phenotype more susceptible to fatigue in the presence of an underlying disease. While this may be an academic exercise at the moment, the existence of a "fatigue" phenotype or a common fatigue pathway paves the way for the thought that potentially,

people more susceptible to fatigue could be identified and the symptom amenable to the same intervention trialled in different disease settings.

In rheumatoid arthritis, fatigue was rated as an essential domain to be absent or reduced/improved to define remission from the patient's perspective (Harris, 2001), and it is a recommended core outcome measure in RA clinical trials (Nikolaus et al., 2013, Minnock et al., 2018, Van Tuyl et al., 2015). Vitality, a proxy for the absence of fatigue, is one of the patient's preferred outcomes of RA treatment(Van der Elst et al., 2020).

2.2.4.3 Measuring fatigue

Measuring fatigue can be challenging due to the lack of a universal definition and the subjective and multidimensional nature of the construct. Fatigue has been measured by several tools across multiple disease conditions; some tools are disease-specific while others are generic, relating to quality of life and general well-being.

Recently attempts have been made to identify biomarkers to objectively measure fatigue in disease conditions. Structural and functional magnetic resonance imaging (MRI) studies to identify brain correlates of fatigue across several chronic disease conditions associated with fatigue have so far met with limited success (Goñi et al., 2018).

In clinical practice and research, fatigue is widely measured using patient-reported outcome measures (PROMs). These PROMS vary in focus, some being unidimensional, for example, relating to severity only, while others are multidimensional, capturing information not only on

severity but also on consequence and impact. Recall periods for fatigue measurement scales also vary, ranging from one day to one year .

An ideal measure of fatigue should be easily understandable, time-efficient, and have good discriminant ability. The tool should capture the multiple dimensions of fatigue and have excellent psychometric properties in terms of reliability, validity, and sensitivity to change(Whitehead, 2009).

A 2007 systematic review identified 23 scales used to measure fatigue. These scales ranged from a binary (yes/No) measure of fatigue to a 15-item multidimensional measure. Only six of these scales had evidence of "reasonable" validation for measuring fatigue in RA. These include the Multidimensional Assessment of Fatigue scale (MAF), Functional Assessment of Chronic Illness Therapy fatigue scale (FACIT-F), Ordinal scales, Profile of Mood States (POMS) fatigue/inertia, Short Form 36 Health Survey (SF36) vitality subscale and the fatigue Visual Analogue scale (VAS) (Hewlett et al., 2007). More recently, the Patient-Reported Outcome Measurement Information System (PROMIS®) has been used as a fatigue measurement scale, but there is limited evidence of its responsiveness (Hewlett et al., 2007, Bartlett et al., 2015). The Bristol Rheumatoid Arthritis Fatigue Multidimensional Questionnaires (BRAF-MDQ), and the BRAF-Numerical Rating Scale (NRS) are the latest addition to the list of fatigue measuring instruments. The BRAF-MDQ was developed to measure a multidimensional fatigue experience. BRAF-MDQ was abridged to assess fatigue severity, effect and coping with fatigue in the BRAF-NRS (Hewlett et al., 2018, Hewlett et al., 2011). Both

scales measure fatigue over the past week. These questionnaires were developed and validated in RA populations and showed excellent internal consistency (Cronbach's α = 0.93), test-retest reliability (r = 0.95 and r = 0.92 for BRF-NRS fatigue severity) for total fatigue.

Other widely used fatigue scales include the fatigue VAS and SF36 vitality scales. Due to its extensive usage, the VAS scale is discussed alongside the SF36 vitality subscale (which was employed in this thesis) in this section.

SF36 Vitality (Vitality) subscale is a four-item tool (full of pep, energy, worn out, tired) with six responses (all of the time - none of the time (Schafer and Graham, 2002, Ware and Sherbourne, 1992, Ware Jr, 1999). This tool measures vitality as a proxy for fatigue, with low scores equating to feeling tired and worn out all of the time and high scores to feeling full of pep and energy all of the time. This tool is based on the concept that fatigue represents a continuum, with vitality at one end of the spectrum and fatigue at the other end. Two versions of SF36 vitality were published in 1992 (SF36) (Ware and Sherbourne, 1992) and 1993 (SF36 version 2)(McHorney et al., 1993). The vitality subscale of SF36 contains six items, while SF36 vitality version 2 contains five items. In RA, SF36 vitality subscale demonstrated good internal consistency (α = 0.84 to 0.88 at multiple time points) when compared with other fatigue scales)(Bacalao et al., 2017). SF36 vitality exhibited excellent test-retest reliability in an RA population (interclass correlation coefficient (ICC) = 0.91)(Linde et al., 2008) construct validity and sensitivity to change in different populations of people with RA; however, there is some evidence that it is strongly

correlated with mood which may suggest a poor differentiation between fatigue and depression (Carandang et al., 2016). In addition, there is no standardised time frame captured by the question; for example, some studies ask about the past week, while others ask about the past month. This makes comparing results from different studies a challenge. A shorter version of SF36, SF12 was published in 1996(Ware et al., 1996) and includes only one item assessing energy.

The fatigue visual analogue scale (VAS) is a 100mm horizontal line with written descriptions at each end; subjects are asked to mark a point on the line that best describes their perception of their current state; with possible scores ranging from 0 (no fatigue) – 100 (very severe fatigue)(Smith et al., 2011). The fatigue VAS has demonstrated good face and content validity, and 'reasonable' construct validity. Unlike the SF36 vitality, VAS is only moderately associated with mood but strongly correlated with axial pain. VAS is more sensitive to change than SF36 vitality; however, lack of standardization limits its use (Hewlett et al., 2007, Tseng et al., 2010). There is evidence of a high correlation among scales measuring fatigue severity and impact. However, it is pertinent to note that the time frame referenced by these scales varies; and the impact of this variability has not been studied in RA patients (Giannouli, 2017).

2.2.4.4 Mechanisms of fatigue in RA

The exact causal mechanism of RA fatigue remains unclear. This is possibly due to the complex nature of fatigue and the lack of standardised fatigue measures making comparisons among studies challenging.

Broadly, however, several factors such as disease activity, pain, mental health status and comorbidities are influenced by fatigue.

Several models explaining fatigue in RA have been described in the literature. A conceptual model of fatigue in RA was proposed by Hewlett et al. in 2010 and has been widely quoted in the literature (Hewlett et al., 2010). In this model (shown in figure 2.3 below), three main contributing factors were identified - RA disease activity encompassing inflammation, joint damage, medication and disability, cognitive and behavioural factors which include illness beliefs and stress, and emotions and finally, personal factors, which include the social environment of the individual, as well as general health. Although popular, scientific evidence supporting this model is mainly cross-sectional and conflicting (Nikolaus et al., 2013). That being said, some of the evidence underpinning some aspects of the model is discussed below.

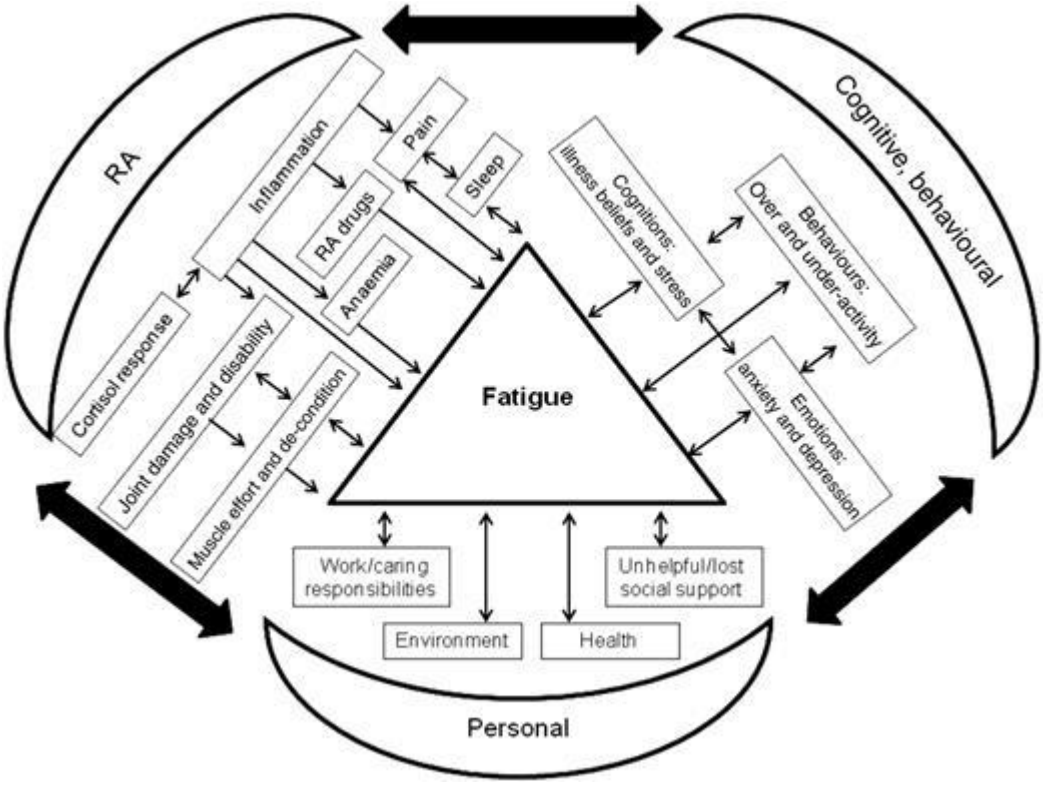


Figure 2.3: Conceptual model of fatigue in RA(Hewlett et al., 2010)

2.2.4.5 Inflammation

Fatigue is thought to be linked with systemic and local effects of the disease. Systemic inflammation, demonstrated by the presence of inflammatory biomarkers, notably Tumour Necrosis Factor (TNF)-α, Interleukin (IL)-6 and C - reactive protein (CRP), have been linked with fatigue. These inflammatory mediators interfere with neurotransmitter and neuroendocrine systems leading to altered behaviour and fatigue. This has been demonstrated clinically in some disease conditions such as breast cancer; however, there is a suggestion that the effect may vary with disease condition(Wang et al., 2017).

Chronic inflammation has also been shown to affect the hypothalamic-pituitary-adrenal (HPA) axis. HPA axis regulates the production of cortisol which in turn suppresses pro-inflammatory response. The presence of

chronic inflammation blunts this response. Cortisol and other glucocorticoids produced by the HPA axis are associated with glucose mobilisation and energy production. Fatigue and lack of motivation are defining features of hypoadrenalism. HPA blunting response has been noted in patients with RA (Crofford et al., 1997). Cortisol and other glucocorticoids produced by the HPA axis are associated with glucose mobilisation and energy production. Fatigue and lack of motivation are defining features of hypoadrenalism. However, the association between HPA axis and fatigue in RA has not been extensively studied.

In patients with RA, it appears that other factors may be involved in the pathogenesis of fatigue. This is evidenced by inconsistent findings from studies on the correlations between markers of disease state; for example, Swollen/tender joint score and ESR were found to be uncorrelated with fatigue, while DAS28 and CRP were found to be correlated with fatigue. Other studies have reported no association between fatigue and inflammation in RA (Giannouli, 2017, Pollard et al., 2006, Repping-Wuts et al., 2007, Van Hoogmoed et al., 2010, Druce et al., 2016). Fatigue has been found to be a persistent symptom in intensively treated patients with low disease activity (Walter et al., 2018). In addition, a systematic review of the effect of biotherapies in RA patients revealed a significant but small reduction in fatigue. It is important to note that participants in these trials had high levels of disease activity on enrolment, and fatigue was a secondary outcome. The inconsistencies surrounding the association between fatigue and inflammation may be due to an indirect link between fatigue and inflammation or other factors

such as pain and depression (Druce et al., 2015a, Katz et al., 2016). It could also be a function of the multidimensionality of fatigue measured in these studies, whereby certain markers of disease factors are associated with specific dimensions of fatigue. Overall, although intuitive, the association between fatigue and inflammation requires further scrutiny.

2.2.4.6 Pain

Pain is a complex multidimensional experience and often a critical feature of rheumatic diseases, particularly RA. Pain is an important predictor of fatigue in RA. In a systematic review, 10 out of 13 studies found a significant association between pain and fatigue. Eight of these studies were cross-sectional. Of the five longitudinal studies included in the review, one found a significant association between pain and fatigue variability; three showed no significant association, and only one demonstrated a significant association between pain and fatigue. The majority of the longitudinal studies discussed had small sample sizes (less than 500), with varying follow-up periods, only one lasting more than two years (Nikolaus et al., 2013).

There is a paucity of literature on the temporal relationship between fatigue and pain in RA. This is probably because RA fatigue and pain were traditionally thought to be consequences of inflammation. Fatigue and pain could be part of the body's adaptive restructuring of priorities in periods of certain stressors, for example, inflammation. Sickness behaviour has been demonstrated in humans and other

vertebrates(Lopes et al., 2016). Pain could be a system to protect the body against impending damage through sickness behaviour.

On the other hand, pain could be an activator of sickness behaviour, leading to a state of increased sensitivity to enable the body to remain hyper-alert against potential stressors(Omdal et al., 2019). A widely quoted longitudinal study reported that changes in pain occurred concurrently with changes in fatigue, suggesting a synchronous relationship, with none preceding the other, alluding to the former in RA. This study was based on a sample of 228 participants with established disease with an overall study duration of one year. This relationship has not been explored in individuals with early disease or over a longer duration (Dartel et al., 2013).

In the conceptual model above, pain in RA is linked to inflammation, but in general, and in RA, pain is a multidimensional and multicausal experience that could result from interactions between multiple mechanisms in the peripheral and central nervous system. The pathophysiology of chronic pain involves both central, peripheral, and cognitive mechanisms as single or overlapping processes. Individuals with RA pain describe it as aching or gnawing, suggesting a nociceptive-type pain, and some describe it as 'shooting' or 'burning', which is characteristic of neuropathic pain (*"Pain caused by a lesion or disease of the somatosensory nervous system"(*International Association for the Study of Pain (IASP)*))*(Walsh and McWilliams, 2014). RA pain may have nociplastic features, described as widespread, and without evidence of

commensurate tissue or nerve damage. The international association on the study of pain (IASP) defines nociplastic pain as

"Pain that arises from altered nociception despite no clear evidence of actual or threatened tissue damage causing the activation of peripheral nociceptors or evidence for disease or lesion of the somatosensory system causing the pain" (International Association for the Study of Pain (IASP))

Increased sensitivity of nerve fibres can occur due to the continuous presence of noxious stimuli, resulting in a reduced nerve activation threshold and a sustained pain experience. This process is called sensitisation. Nociplastic pain may be due to central sensitisation (Chimenti et al., 2018, Roche et al., 2003) and is not caused by inflammation and therefore not amenable to anti-inflammatory agents. Studies have shown persistent pain in patients with low RA disease activity(Repping-Wuts et al., 2007). Druce et al. reported that in a population of people with RA with severe fatigue who failed to respond to conventional DMARD therapy, improvements in fatigue observed following anti-TNF therapy represented improvements in central pain mechanisms (Druce et al., 2016). This study alludes to a relationship between fatigue and nociplastic-like pain and, by extension, central pain sensitisation. Central sensitisation has been demonstrated in RA and other chronic inflammatory conditions characterised by pain such as fibromyalgia and Ankylosing Spondylitis, which are conditions strongly associated with fatigue (Rongen-van Dartel et al., 2014, Vincent et al., 2013).

2.2.4.1.1 Fatigue and central sensitisation (CS)

The concept of central sensitisation was introduced in the preceding section during the discussion of fatigue and pain and is explored further in this section of the thesis.

According to the International Association for the Study of Pain (IASP), CS is defined as

"*Increased responsiveness of nociceptive neurons in the central nervous system to their normal or subthreshold afferent input*(International Association for the Study of Pain (IASP))"

Clinically, central sensitisation is assessed through allodynia and hyperalgesia, defined by IASP as "*pain due to stimulus that does not normally provoke pain*" and " *increased pain from a stimulus that normally provokes pain*" (International Association for the Study of Pain (IASP)), respectively, however, studies have shown that CS might reflect increased responsiveness to other stimuli such as cold, heat, chemical substances, electrical stimulus, stress, mental load, emotions and sensitivity to environmental stimuli such as weather(Nijs et al., 2014).

CS is associated with generalised sensory hypersensitivity, where individuals report sensitivity to other sensory stimuli such as visual, olfactory, and auditory stimuli. This has been demonstrated in conditions significantly associated with fatigue, such as fibromyalgia (Geisser et al., 2008), migraine(Goadsby et al., 2017), Irritable Bowel Syndrome and temporomandibular joint disorder (TMD) (Blomhoff et al., 2000).

General hypersensitivity may be associated with the activation of the anterior cingulate, prefrontal cortex and the insula, which comprise the cortical network, areas responsible for higher-order neural processing. Functional MRI studies have demonstrated an association between abnormal functioning of these brain areas with fatigue in chronic fatigue syndrome and multiple sclerosis (Nijs et al., 2012, Wortinger et al., 2017)

CS in chronic pain conditions has undergone extensive research in the past decade; therefore, most available information about CS is based on findings from research on conditions characterised by chronic pain, e.g., fibromyalgia, osteoarthritis, and low back pain. These studies show that the prevalence and severity of central sensitisation exhibit substantial person-to-person variability in many of these conditions, occurring in almost all patients with fibromyalgia and in a proportion of people with knee OA(Lluch et al., 2014, Cagnie et al., 2014).

Several methods exist to measure CS, including imaging, quantitative sensory testing (QST) - a battery of test modalities aimed at assessing the function of the somatosensory system, e.g., pain, pressure and thermal thresholds, and self-report questionnaires which utilise patient-reported outcome measures (PROMs) incorporating some common symptoms of CS(Williams, 2018, Neblett, 2018).

2.2.4.7 Psychological factors

2.2.4.7.1.1 Affective Disorders

One of the key predictors of fatigue in RA is depression. The exact mechanism is difficult to describe. This is because depression is

associated with other factors related to fatigue, such as low socioeconomic status, pain, poor sleep, and inflammation. Depression may also influence how people report their symptoms (Margaretten et al., 2011). Also, lack of energy is a diagnostic criterion for depression, although not everyone with fatigue experiences depression. In addition, some measures of fatigue for example, SF36 Vitality exhibit poor differentiation between fatigue and depression (Nikolaus et al., 2013, Katz et al., 2016).

2.2.4.7.1.2 Cognitive Styles

Studies have shown that certain cognitive styles such as catastrophizing, low physical and functional capacity and learned helplessness (where individuals respond passively to problems they encounter because they feel they have little control over life events) are associated with more significant fatigue. Conversely, self-efficacy has been found to have an inverse relationship with fatigue(Camacho et al., 2013).

2.2.4.8 Co-morbidity/multimorbidity

Comorbidity describes the burden of one disease co-existing with another disease of interest, and multimorbidity is defined as the simultaneous presence of two or more diseases in the same individual linked through a pathogenetic mechanism (Johnston et al., 2019). Common comorbidities present in RA are cardiovascular diseases, lung cancer, gastrointestinal disorders, depression, osteoporosis, and lymphomas. The presence of comorbidities is associated with greater functional impairment and shortened lifespan. Several studies have

linked the presence of comorbidities with severe fatigue in RA (Dougados, 2016, Turesson, 2016, Tournadre et al., 2018).

2.2.4.9 Sleep Disturbance

Sleep disturbance is prevalent in RA, occurring two to three times more commonly in these patients than in the general population. In RA, the relationship between fatigue and sleep is still not well defined. Poor sleep is associated with increased daytime sleepiness, resulting in fatigue, with daytime fatigue leading to poor sleep (Lewis et al., 2019). Poor sleep may be affected by abnormal synthesis of neuroendocrine sleep mediators such as melatonin, prolactin and growth hormone caused by inflammation. Poor sleep could also lead to a vicious cycle, resulting in changes in the normal circadian pattern of circulatory cortisol, which in turn regulates sleep(Guyon et al., 2014).

Previous studies have demonstrated a link between sleep disturbances and fatigue (Irwin et al., 2012). RA fatigue is associated with several domains of sleep, including sleep quality, efficacy, efficiency and latency, frequency, and times of awakenings with a higher proportion of sleep disorders compared to the general population(Sariyildiz et al., 2014). As mentioned above, sleep quality is affected by a host of factors, including chronic inflammation, pain, and medication. All these factors are associated with fatigue (Kilic et al., 2015, Sariyildiz et al., 2014, Westhovens et al., 2014).

2.2.4.10 Other factors

Obesity, low physical activity, and skeletal muscle weakness are also linked with fatigue. Some studies have reported a direct link between obesity and fatigue; however, there may be an indirect link through sleep disorders caused by obesity. It could also be a consequence of chronic inflammation, which affects body metabolism and leads to loss of lean body mass (rheumatoid cachexia), causing reduced muscle strength and increasing the functional burden of daily activities, presenting as fatigue.

Low physical activity has been directly linked to more fatigue. This may be due to the increased secretion of inflammatory cytokines and/or reduced cardiovascular fitness. Conversely, higher physical activity levels may mitigate rheumatoid cachexia and cause improved sleep, which in turn reduces fatigue(Rongen-van Dartel et al., 2014, Løppenthin et al., 2015).

2.2.5 Treatment of fatigue

The causal pathways of fatigue in RA are poorly understood, and as a consequence, there is no gold standard management protocol for fatigue in people with fatigue in RA. Pharmacological and non–pharmacological interventions have been tested, with findings summarised below.

2.2.5.1 Non-pharmacological

A systematic review identified several interventions used to manage fatigue in RA. These could be broadly grouped into physical activity, psychosocial and other interventions. Other interventions include a

Mediterranean diet, herbal medicine, reflexology, and omega-three fatty acid usage.

Meta-analysis of six moderate to low-quality randomized control trials (RCTs) showed that physical activity interventions were more effective in the intervention group than in the control group, although with a small effect size(Cramp et al., 2013).

A small RCT focusing on a person-centred physical activity intervention improved fatigue levels in the intervention group compared to the reference group(Feldthusen et al., 2016a). Another RCT found an improvement in sleep quality and fatigue after a twelve-week targeted physical activity intervention(Durcan et al., 2014). A study using daily step diaries, pedometers, and education found improvements in fatigue levels after 21 weeks of follow-up in the intervention groups compared to the education-only group(Katz et al., 2018). These were small RCTs with short follow-up periods and without standardised physical activity programmes.

A similar small effect was found with psychosocial interventions such as cognitive behavioural therapy (CBT), expressive writing, and education. Other interventions such as a Mediterranean diet, reflexology and omega-three fatty acid supplementation, appeared to show statistically significant improvement over controls. However, the quality of these studies was rated as low(Cramp et al., 2013). Overall, CBT showed the most promise, even though some results are conflicting. In some studies, CBT improved fatigue and pain (Thomsen et al., 2017, Hewlett et al.,

2019b). However, another study using an internet-based cognitive-behavioural intervention found little or no difference in participants' pain and fatigue after 12 months of follow-up (Ferwerda et al., 2017). Additionally, more research is required to provide cost-effective psychological interventions to manage fatigue in RA(Hewlett et al., 2019a, Ferwerda et al., 2018).

2.2.5.2 Pharmacological

A recent Cochrane meta-analysis of biologic DMARD therapy on fatigue, including 32 studies and over fourteen thousand people with RA, showed a small but significant reduction in fatigue, with a standardized mean difference (SMD) of -0.43 (CI: -0.49 to -0.38). Other more recent trials showed similar effects (Strand et al., 2016a, Strand et al., 2018a, Strand et al., 2017b). Additionally, clinical trials involving targeted synthetic DMARDs e.g tofacitinib and baricitinib showed a significant reduction in fatigue levels compared with placebo or conventional DMARDs (Strand et al., 2017a, Strand et al., 2018b, Strand et al., 2015, Emery et al., 2017, Schiff et al., 2017, Keystone et al., 2017, Smolen et al., 2017). Most of these trials were of moderate quality, and fatigue was measured with various instruments. The trials were set up to examine the efficacy of the medications on RA disease activity and fatigue was a secondary outcome. It is also unclear if the improvements in fatigue scores were a direct result of medication or the reduction in disease activity or pain (Almeida et al., 2016, Choy, 2019).

Fatigue remains a persistent problem despite breakthroughs in RA pharmacotherapy and more intensive treatment regimens. Non-pharmacological therapies have also shown limited success. The scientific community agrees that there are multiple causal pathways to fatigue, but less explored is the thought that these pathways are not generic to all individuals, and maybe some individuals may have a dominant pathway. Nikolaus et al. demonstrated four different patterns of fatigue from the patient's perspective, ranging from little impact of fatigue to high distress. The study concluded that fatigue is influenced by disease-specific as well as psychological and social aspects (Nikolaus et al., 2010). The study was conducted on a small sample (30 participants), and data were collected at a one-time point; it is not entirely clear if these patterns are sustained over time, but it may be that the patterns described represent different dominant causal pathways. It can therefore be inferred that people who receive treatment targeted at their dominant pathways would exhibit less disability and other adverse effects of fatigue.

2.2.6 Summary of literature review

In summary, fatigue constitutes a significant burden to patients, their families, and the society at large. It is one of the most distressing symptoms people with RA must deal with, and these individuals often feel the symptom is ignored by health care providers. There is no unified definition of the construct, and no objective measure of fatigue. Consequently, it is difficult to estimate the actual prevalence of the symptom in people with RA, as well as in other disease conditions. In RA,

several putative factors such as pain, disease activity and inflammation, co-morbidities, sleep, and other psychosocial factors have been suggested to be associated with or play a causal role in the development of fatigue.

Nevertheless, evidence is conflicting; hence, fatigue causal mechanisms are poorly understood. Perhaps, consequently, both pharmacological and non-pharmacological interventions documented to treat fatigue in RA have shown minimal effects at best and uncertain long-term effects. Historically, in RA, fatigue research has focused on people with long-standing and established disease, with scarce data from people with early disease. Not everyone with RA suffers from fatigue. However, the natural history of fatigue has not been clearly described. In addition, it is unclear if inter-individual differences exist in people with fatigue. Identifying a target population more likely to experience fatigue in RA early in the disease course could potentially provide insights into the natural history of fatigue and represent an ideal population for targeted interventions to alleviate fatigue in RA.

2.3 Study rationale

Fatigue is a significant symptom in RA patients. Its exact cause is unknown, although it is thought to be at least partially explained by inflammation. However, a lot of the available evidence is based on cross-sectional data. There is no standardised way to measure fatigue, and there is a paucity of good-quality longitudinal data on the relationship between fatigue and other important variables in RA.

Patients with RA fatigue may represent a heterogeneous population, and identifying these sub-populations may provide further insight into the causal mechanisms of RA fatigue.

Another critical point to consider is that majority of epidemiological studies on fatigue in RA are based on individuals with established disease. Some studies allude to central mechanisms in the pathogenesis of fatigue similar to pain sensitisation; however, studying individuals with established disease who may be already sensitized may mask the true nature of fatigue in these individuals.

Characteristics associated with fatigue bisect many unrelated disease conditions. Perhaps certain individuals possess traits which increase susceptibility to fatigue in the presence of an underlying disease.

This project aims to unpick some of the theorized associations between fatigue and covariates in order to gain a better understanding of fatigue in rheumatoid arthritis. I aim to shed light on the longitudinal relationship between fatigue and important covariates, test the theory that there may be groups of people who share similar fatigue trajectories and attempt to identify factors associated with these fatigue sub-populations.

2.4 Aim and objectives.

2.4.1 Aim

The overall aim of this thesis is to increase the understanding of fatigue in RA.

2.4.2 Objectives

The objectives of the studies are as follows:

1. To determine the prevalence and characteristics associated with fatigue in RA.
2. To describe the longitudinal course of fatigue and examine heterogeneity in the course of fatigue in RA.
3. To describe the prevalence and associations of fatigue in the general population.
4. To design an observational study to examine the association between fatigue and central pain mechanisms in RA.

2.4.2.1 Hypothesis

The hypothesis governing this study is that specific individuals may be more susceptible to the development of fatigue in RA, and these individuals could be identified early in the disease course.

2.4.2.2 Study outcomes

This study will add to the body of evidence on associations between fatigue and important covariates, such as pain, depression, comorbidity, medication, markers of disease activity and quality of life.

Finally, my proposed hypothesis could facilitate more focused research and ultimately lead to the design of targeted interventions and hopefully better outcomes for individuals who suffer from fatigue in RA and other disease conditions.

3. Chapter 3 Introduction to methods

3.1 Outline

This chapter describes the study populations, main outcome variables, covariates, and strengths and limitations of each dataset used in statistical analyses conducted in chapters four, five and six. This chapter also provides an overview of the statistical methods used to deliver results in those chapters. Brief descriptions of the methods used in each chapter are discussed in respective chapters to ease readability. Finally, this chapter is divided into two broad sections. The first section describes the study datasets, and the second, the analytical methodology.

3.2 Study datasets

3.2.1 The early rheumatoid arthritis network (ERAN)

ERAN is an inception cohort of rheumatoid arthritis (RA) designed by the ERAN registered charity. Participants were recruited from 23 rheumatology departments in England, Wales, and Ireland between 2002 and 2012. Geographical locations included rural, urban and ethnically diverse were selected to represent the UK population as much as possible(Garwood, 2004).

At baseline, participants with a clinician's diagnosis of rheumatoid arthritis who consented to participate in the ERAN study were enrolled if they had a disease duration of fewer than 24 months. Participants were excluded if found to have an alternative diagnosis at subsequent visits and therefore not analysed in this study. ERAN contains a wealth of

information on early RA and has been used in several published articles on the epidemiology of early RA and in the development of NICE decision rules(Gibson et al., 2015).

3.2.1.1 Ethics

ERAN received ethical approval from the UK National Health Service, Trent research ethics committee reference 01/4/047.

3.2.1.2 Study design

Data were collected from ERAN participants at baseline (the first measurement occasion), three to six months later (the second measurement occasion), and annually after that. For this thesis, data from the first five measurement occasions were used because of the high attrition rates at later study measurement occasions.

3.2.1.3 Measures

3.2.1.3.1.1 Outcome variable

Fatigue was measured using SF36 Vitality (Vitality). Vitality has been widely used to assess fatigue in many disease conditions, including rheumatoid arthritis (Linde et al., 2008). It exhibits good psychometric properties in people with RA across different geographic distributions (Loge et al., 1998b, Feroz et al., 2012, Ruta et al., 1998).

3.2.1.3.1.2 Data management and interpretation of Vitality

According to the developer's instructions, the four Vitality items were transformed to a 0 – 100 scale. To enable comparison with the general population of respective geographical settings, in this case, the UK, transformed Vitality was normed to the average UK population, as

recommended by the developers (Ware, 1993). Vitality is inverse scored such that higher values represent a better quality of life (less fatigue). Additionally, clinically significant Vitality was interpreted based on the recommended minimum important difference (MID) of Vitality throughout this thesis as a 10-point change/difference(Bjorner et al., 2007).

3.2.1.4 Exposure variables

Demographic characteristics: self-reported age at onset of disease (in years), sex (male/female), body mass index (BMI) in kg/m^2, and self-reported smoking status (never, current, and ex-smoker).

Inflammation and disease activity: ESR was used as the main inflammation index and measured on a continuous scale in mm/hr.

Tender (TJC) and swollen joint counts (SJC) in the proximal interphalangeal (PIP), metacarpophalangeal (MCP), wrist, elbow, shoulder, and knee joints were reported in ERAN. TJC and SJC were measured as a count from 0 – 28.

Patient's Global Assessment of Disease Activity (PGA) was reported on a visual analogue scale (VAS) of 0 – 100mm, where higher values represented greater impact(Nikiphorou et al., 2016).

DAS28ESR was used to measure disease activity in ERAN. In this thesis, component indices of DAS28 were used in multivariable analysis.

Pain and mental health were measured with SF36 bodily pain subscale (Pain) and SF36 mental health subscale (mental health), respectively. Pain measures the presence and impact of pain and combines responses

from two questions while mental health assesses mental symptoms and psychological wellbeing by combining responses to five questions, all covering a period of four weeks. Mental health is an umbrella term used to describe diverse conditions that are managed in different ways. A granular description of the impact of mental health on fatigue was therefore not possible with this dataset.

Like Vitality, these subscales were normalised to the UK population on a 0 – 100 scale, with higher values representing a better QoL (less pain and better mental health). Scoring of Vitality, pain and mental health was done using a STATA® scoring script developed by Philip Ryan, and adapted to the UK population by Dr Samuel Norton(Ryan, 1999).

Disability/Functional Limitation was measured using the health assessment questionnaire (HAQ) disability index (HAQ-DI). This questionnaire assesses the level of functional ability by examining functional activities such as walking, eating, reach, grip, and usual activities. It is scored on a scale of 0 – 3, with higher scores indicating greater disability(Wu et al., 2021).

Comorbidities were assessed using the rheumatic disease comorbidity index (RDCI). The RDCI was developed to quantify the total burden of comorbidities in people with RA. It was created from self-reported questionnaires from patients with musculoskeletal and inflammatory diseases including RA. RDCI is a validated measure of comorbidities in RA. It is calculated using the formula in equation 3.1. In ERAN, component indices of the RDCI were obtained from text entries of

component disease conditions, with total scores on a scale ranging from 0 – 9. Higher scores represent a larger comorbidity burden(England et al., 2015).

Equation 3.1: Composition of RDCI

$$RDCI = 2 \times lung\ disease + \lfloor (2 \times heart\ attack, other\ cardiovascular\ disease, OR\ stroke)\ OR\ 1 \times hypertension \rfloor + fracture + depression + diabetes + cancer + ulcer\ or\ stomach\ problem$$

Disease duration: was measured by the self-reported time since onset of symptoms in months.

Medication use was represented by a binary (y/n) variable, with the affirmative comprising participants prescribed disease-modifying anti-rheumatic drugs (DMARDs).

Haemoglobin concentration was measured at each measurement occasion in grammes/decilitre (g/dl).

Time was assessed as the study measurement occasion. Data were used up to the fifth measurement occasion (three years from baseline).

Other covariates included rheumatoid factor (RF), and anti-cyclic citrullinated peptide antibodies (anti-CCP). These markers were used to derive a variable depicting seropositive RA. This variable called seropositivity was defined by the presence of RF and/or positive or borderline anti-cyclic citrullinated peptide antibodies (Lee et al., 2008). The presence or absence of nodules on clinical examination were also assessed.

3.2.1.5 Strengths and Limitations of ERAN

ERAN is an inception cohort with many patients and repeated measurements, allowing the assessment of longitudinal and cross-sectional associations. ERAN included demographic, clinical, laboratory and patient-reported outcomes in a nationally representative sample.

In addition to the limitations of covariates collected in the ERAN data described above, high attrition rates, especially at later stages of the study, led to the closure of the cohort. In this thesis, the inclusion of only five measurement occasions precluded the use of some more complex statistical modelling of longitudinal data.

3.2.2 Knee pain and related health in the community (KPIC) study

The KPIC study is a general population cohort study designed to determine the prevalence and variation of self-reported knee pain characteristics in the community. This study was conceptualised and implemented by the academic rheumatology department of the University of Nottingham(Fernandes et al., 2017).

3.2.2.1 Ethics

Ethical approval was obtained from the Nottingham Research Ethics Committee 1 (NREC Ref: 14/EM/0015).

3.2.2.2 Study setting and design.

Postal questionnaires were sent to adults aged 40 and above who were registered at general practice (GP) surgeries in Nottinghamshire, East Midlands. GPs were approached through the East midlands clinical research practice network and surgeries willing to participate were

recruited. Eligibility to take part in the KPIC study was determined by the healthcare professional from a population of people included in the GP register in respective GP surgeries. The GP register is a log of individuals who have lived in the local area for a minimum of 24 hours and are registered with the local National Health (NHS) service. 40,000 questionnaires sent to participants contained questions about general health, in addition to knee pain-related questions. 9,513 questionnaires were returned. Approximately one year later, a second questionnaire was posted to 6716 participants who indicated a willingness to continue to take part in the study. 4,738 questionnaires were returned. Figure 3.1 shows a flow chart of KPIC recruitment (Fernandes et al., 2017). Participants were excluded from the study if terminally ill, had major mental illness, dementia, or any other illness, making them unable to receive or complete the questionnaire. The general practitioners (GPs) made this decision. Data collected at baseline and year one were used in this thesis.

A sample of baseline responders who indicated a willingness to consider having knee radiographs were identified based on responses to questions about the duration and severity of knee pain. Two hundred participants with early-onset knee pain, defined as mild/moderate and or intermittent knee pain occurring for the first time in the past three years, occurring on most days for at least one month, and unrelated to major trauma were identified. These participants were age and sex-matched with the established knee pain (n=100) and no-knee pain (n=100) groups. Established knee pain was defined as moderate knee pain, measured on

a numerical rating scale > 6, on a scale of 0 – 10, and/or occurring on most days of the past three months unrelated to major trauma. No-knee pain was defined as the absence of knee pain within the last five years.

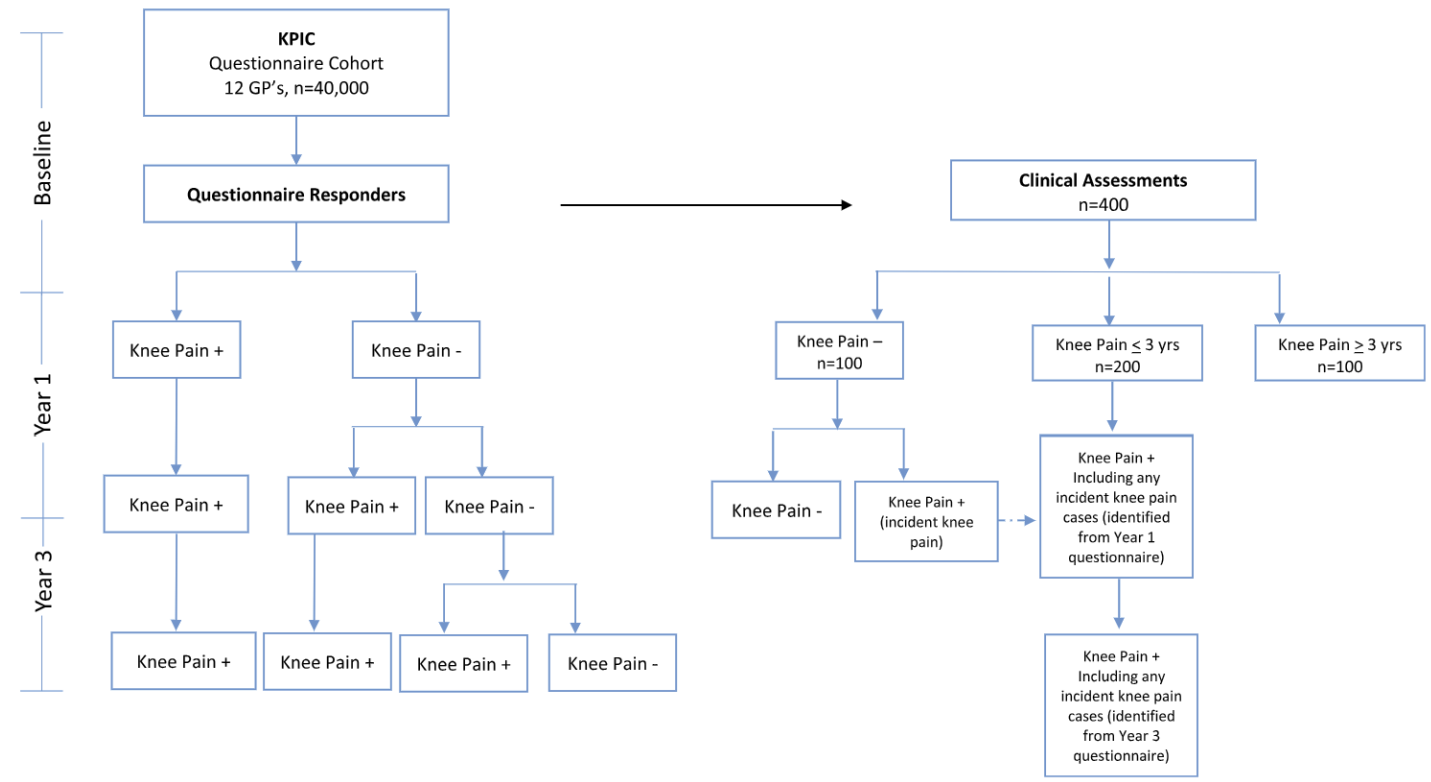


Figure 3.1:KPIC study recruitment flowchart ,source:(Fernandes et al., 2017)

Participants were excluded if pregnant, reported major knee trauma, or had total knee replacement surgery. Participants recruited for further clinical assessments were invited to an assessment visit, where blood samples were collected for markers of inflammation, and a number of clinical assessments, including pressure pain detection threshold (PPT) were conducted. KPIC data were provided by the KPIC committee (Appendix 11.1.2 KPIC study data access request). Data used in this thesis were analysed according to the sampling framework.

3.2.2.3 Measured covariates and data management

3.2.2.3.1.1 Outcome variable

In KPIC, fatigue was measured using the vitality subscale of the medical outcome study (MOS) 12 Item short-form health survey short-form 12, SF12VT. SF12 was developed from SF36 and covers the same eight domains but with fewer questions. SF12 has been used by policymakers to screen populations and evaluate healthcare utilisation and outcomes. It has been validated in several studies, and the correlation between summary scores of SF12 and SF36 is well documented in many health conditions and in the general population (Wukich et al., 2016, Jenkinson et al., 1997, Gandek et al., 1998).

SF12VT is made up of one item derived from the question *"Did you have a lot of energy?"* and scored on a scale of 0 – 100 as described by John Ware(Ware et al., 1996). Normative values based on the UK population were calculated, as described above in the ERAN cohort. SF12VT is reverse-scored such that higher values represent a better quality of life

(less fatigue), and in this thesis, clinically significant SF12VT was interpreted in the same manner as SF36 Vitality in the ERAN dataset.

In addition to SF12VT, another variable representing fatigue was derived from the KPIC cohort from the following question: "*In the past month, did you feel tired on most days?*" and with responses on a five-point Likert scale – never, seldom, often, sometimes, always. People who answered 'often', 'sometimes' and 'always' were assigned to the fatigue group and those who responded 'never' and 'seldom' in the no-fatigue group.

3.2.2.3.1.2 Exposures

Demographic characteristics recorded included self-reported age (in years), sex (male/female) and body mass index (kg/m^2).

Mental health was assessed using the hospital anxiety and depression scale (HADs). HADs was developed to detect states of anxiety and depression in a general medical population. It is made up of two subscales with seven questions each that assess anxiety and depression, respectively (Wu et al., 2021). HADs has been validated in many languages, in primary care and community settings. It is one of the NICE recommended tools for assessing depression and anxiety (Stern, 2014). HADs subscales are scored on a 0 – 21 scale, with higher values signifying worse outcomes. Cut-off points of between seven and eight have been described in the literature (Stern, 2014). HADs subscales were analysed as continuous variables in this thesis.

Sleep was assessed using the medical outcomes survey sleep scale (MOS-Sleep). MOS-Sleep is made up of seven subscales assessing

several domains, including sleep disturbance and adequacy, somnolence, snoring, sleep short of breath or headache, and sleep problems index I & II. For the purposes of this thesis, the sleep variable was derived from the sleep adequacy subscale generated from one item of MOS-sleep and according to the developer's instructions. Sleep adequacy was scored on a 0 – 100 scale, where higher values represent a better quality of life (better sleep) (Viala-Danten et al., 2008).

Pain was assessed with the bodily pain subscale SF12, scored, and interpreted similarly to SF12VT, according to the developer's instructions.

Inflammation was measured by C-Reactive Protein (CRP).

Comorbidities were measured as a count of self-reported comorbidities. There is no gold standard comorbidity measurement tool. Existing tools are heterogenous, with variations in the range of encompassed disease conditions, weights assigned to specific diseases, and the disease conditions in which the tools have undergone validation (Diederichs et al., 2011, Aslam and Khan, 2018). In rheumatic diseases, disease counts are a common measure of comorbidities and are appropriate in the absence of validated measures (Alyson et al., 2012, Johnston et al., 2019).

Pressure Pain Detection Threshold (PPT) was used as the proxy measure of central sensitisation. PPT is the minimum force required to elicit pain in an individual, and it is one of the threshold modality paradigms that constitute quantitative sensory testing QST(Maquet et al., 2004). PPT taken from the tibia has shown excellent inter and intra-rater reliability in healthy volunteers (Pelfort et al., 2015).

PPT in the anterior tibia, distal to the most painful knee, was used(Akin-Akinyosoye et al., 2018). In the no-knee pain subgroup, PPT obtained from both the right and left anterior tibiae were used separately. PPT was also measured in the sternum to serve as the distal body part(Krüger et al., 2016, Soni et al., 2013).

The measure of central sensitisation using KPIC data is limited because only one modality of QST was assessed. PPT may be influenced by factors outside central sensitisation, such as the events occurring during data collection, e.g., participant attention, reporting styles, the interaction between researcher and participant, and peripheral sensitisation.

PPT is associated with knee pain; however, this association may be due to a peripheral rather than central sensitisation(Mutlu and Ozdincler, 2015). In addition, the closer the site of PPT assessment to the most painful knee, the higher the risk of peripheral sensitisation. It is important to note that only pain sensitivity is assessed in this study, and it is not clear if the response to PPT captures other hypersensitivities.

3.2.2.3.1.3 Strengths and Limitations of KPIC

KPIC is a community-based study with a large sample size designed to represent the general population of the specific study setting (East Midlands). KPIC is a cohort study, with data collected on multiple measurement occasions. This study design facilitates the assessments of multiple outcomes and exposures, including studies not related to knee pain. KPIC contains a broad range of covariates, including radiological,

laboratory and validated patient-reported outcome measures allowing for the assessments of many putative factors associated with fatigue.

KPIC data were collected from the East Midlands only, however back comparison to the UK population using census data from the office of national statistics showed similarities in key demographic variables. KPIC is a questionnaire-based study, and participation was initiated by the GPs. This means that only people who engage with GP services can be contacted for participation. Data from office of national statistics (ONS) show that the population of individuals in the GP register is representative of the UK population. In addition, considering the reported proportion of people with bad or very bad health the in Health Survey for England (HSE), and assuming an average number of registered patients in each participating GP practice, 40,000 roughly approximates the total number of people aged 40 and above in the GP register(Conolly and Davies, 2018, NHS Digital, 2017).

Furthermore, participants with knee pain are more likely to respond, creating a bias towards people with knee pain through response bias. Additionally, CRP and PPT measurements were taken in a fraction of participants at only one time point and therefore not included in the longitudinal analysis.

The lack of a validated comorbidity variable within the KPIC dataset represents an additional limitation. Nevertheless, utilizing comorbidity counts serves as an acceptable alternative for measuring comorbidities in instances where a validated comorbidity index is unavailable.

3.3 Statistical Analyses

3.3.1 Direct standardisation

Comparing health indices such as incidence or prevalence rates of diseases between geographical settings, in population subgroups or at different time points in the same population is fundamental to epidemiological studies. It facilitates the identification of risks, needs definition, and helps highlight health inequalities if present. Large populations are heterogeneous in terms of age, gender, education, and religion; therefore, directly estimating these rates (crude rates) can be misleading if this diversity is not considered.

Standardisation is a method used to compare rates by eliminating the influence of extraneous factors(Anderson, 1998, Burack et al., 1983). There are two main standardisation methods, direct and indirect standardisation. Direct standardisation was used in this thesis to describe the prevalence of fatigue at baseline in RA and in the general population and will be briefly described in the following paragraphs.

Direct standardisation estimates the weighted average of group-specific rates, where weights are taken from a standard population. A standard population is a theoretical population, defined as having a particular demographic distribution. There are several standard populations, including the World Standard Population, the standard population of OECD countries and the European Standard Population (ESP). For the purpose of this thesis, the 2013 ESP was used as it is better suited for standardisation in European countries.

Direct standardisation is calculated by stratifying the population of interest into groups based on the variables to which the population will be standardised (in this case, age, and sex). Group-specific rates are estimated and applied to the corresponding group in the standard population. These group-specific rates are then summed to give the overall total population rate.

Equation 3.2: Direct standardisation formula

$$R_s = \sum (r_k W_k)$$

Where:

r_k = rate in the kth stratum of the study population

N_k = number of persons in k^{th} stratum of the standard population

N = total number of persons in the standard population (ΣN_k)

Wk = weight for each stratum (equal to Nk/N)

$\sum$ = summation over the k strata

R_s = Standardised rates

Although standardisation provides a means to compare rates across different populations, it is not without limitations. Standardised rates are average measures obtained from an arbitrary 'standard' population; therefore, may not indicate the true population rates. Standardised rates may not be comparable to rates obtained using different standard populations (Schoenbach and Rosamond, 2000).

3.3.2 Correlation and linear regression

Correlation denotes the association between two continuous variables. It measures the strength of a linear association and is represented by the correlation coefficient (r). Values of 'r' are measured on a scale ranging from -1 through 0 to +1, where 0 denotes no association, -1, a perfect negative correlation (one variable increases while the other decreases) and +1 a perfect positive correlation (both variables increase). Correlations can be represented graphically using scatterplots (a two-dimensional representation of measurements of two variables on the x and y axis, respectively) (Asuero et al., 2006). There are several zero-order correlation statistics (correlation with two variables that do not require adjustment for other variables); however, in this thesis, Pearson's product-moment correlation coefficient (equation 4.1) and Spearman's correlation coefficient were used. Pearson's product-moment correlation coefficient was used in the presence of normally distributed data, and Spearman's for non-normally distributed data.

Equation 3.3: Pearson correlation coefficient formula

$$r = \frac{\mathrm{n}(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2]\left[n\sum y^2 - (\sum y)^2\right]}}$$

Where:

r = Pearson corelation coefficient

n= number of the pairs

∑xy = sum of products of the pairs

∑x = sum of the x

∑y= sum of the y

$\sum x^2$ = sum of the squared x

$\sum y^2$ = sum of the squared y

When the statistical significance of zero-order correlations are tested within a study, it is important to adjust for inflated type 1 errors (α - caused when the true null hypothesis is rejected). This problem is handled by estimating the adjusted α value through techniques such as Bonferroni correction(Haynes, 2013)

3.3.3 Linear regression

Linear regression is a statistical modelling technique that uses a straight line to describe the relationship between two variables. Linear regression assumes that this relationship is a linear function, determined by an intercept and a slope. Simple linear regression can be represented by the following equation:

Equation 3.4: Simple linear regression

$$y = a + bx + e$$

Where*:*

y = outcome (dependent) variable

a = intercept

b = slope

e = error term

x = exposure (independent) variable

The intercept represents the value of the dependent variable when the independent variable is 0, and the slope is the increase in the dependent

variable for one unit increase in the value of the independent variable. The error term represents values of the dependent variable that cannot be explained by the relationship between the dependent and independent variables. In linear regression, the line of best fit minimises the errors in the model. This can be achieved by the ordinary least squares (OLS) method.

The amount of error between the value predicted by the regression line and the observed value of y in the equation can also be called the residual. The residual sum of squares (SS) is the sum of each residual squared in the model. This residual sum of squares and the model sum of squares, the variability of the outcome variable explained by the independent variable, make up the total sum of squares. The R^2 statistic is the ratio of the model sum of squares to the total sum of squares and explains the variability of the dependent variable explained by the independent variable. R^2 tell us how well the line fits the data; values range from 0 – 1, with higher values depicting a better fit.

In order to estimate unbiased results, OLS assumes that the mean of the error terms is 0, errors are independent, with equal variances and follow a normal distribution. Simple linear regression includes one independent variable, and multiple linear regression includes more than one independent variable.

In multiple linear regression, the relationship between one dependent variable and multiple independent variables is modelled. It is an

extension of the simple linear regression model and can be represented by the following equation.

Equation 3.5: Multiple linear regression equation

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_k x_k$$

Where:

y = Dependent variable

β_0 = Intercept

$\beta_1 \ldots \beta_k$ = Slope of variable 1 – k, after adjusting for other independent variables

$x_1 \ldots x_k$ = Independent variables 1 to k

Model selection is the process of selecting the optimal number of independent variables in a linear regression model. Model selection attempts to balance simplicity, scientific relevance, and statistical accuracy. In this thesis, stepwise backwards selection modelling technique was adopted for regression modelling(Chowdhury and Turin, 2020). In backward selection, all independent variables are included in the regression model initially, and independent variables are dropped based on statistical significance until all remaining variables are statistically significant. The stepwise method is easy to explain and not computationally intensive; however, optimal models may be missed when variables are added and dropped in turn. This method may overstate significance values because of issues with multiple testing.

The relative strength of associations was an important measure in this thesis; therefore, in addition to the regression coefficient, standardised regression coefficients were estimated. The standardised regression

coefficient was obtained by transforming all independent variables in the final model to z scores before running the regression model. In this case, the independent variables are interpreted as the effect of one standard deviation increase in the independent variable on the dependent variable.

Following model selection, linear regression assumptions were checked. These assumptions include a linear relationship between the dependent and independent variable, usually examined with scatterplots; the dependent variable, y, should be normally distributed for each value of the independent variable, x. This was tested by assessing if the residuals follow a normal distribution. In a linear regression model, the standard deviation of y is the same for each value of x, tested by plotting the residuals against the independent variable and checking that the spread is constant across the range of values. Violation of this assumption was addressed by model estimation using robust standard errors.

Multicollinearity describes a perfect linear relationship among independent variables in a linear regression model. Multicollinearity causes unstable regression coefficients and inflates standard errors. The variance inflation factor (VIF) measures the degree of multicollinearity in a regression model and was used in this thesis. As a rule of thumb, VIF values less than ten were considered acceptable.

3.3.4 Logistic Regression

Logistic regression is a method used to describe the relationship between a binary outcome variable (such as mortality) and single (or multiple) exposure variables. Binary outcomes usually take the form of indicator

variables, represented with numbers 1, indicating the presence of the outcome (e.g., death), and 0, depicting the absence of the outcome variable (e.g., absence of death - alive). This distinguishes logistic from linear regression. This difference affects regression modelling and assumptions. Beyond this, logistic regression follows the same general principles of linear regression described above (Hosmer et al., 2013) (Stoltzfus, 2011).

The logistic regression model estimates the odds ratio of an outcome, given the exposure variable(s). The odds of an event represent the probability of that event happening divided by the probability of the event not happening, while the odds ratio is the ratio of odds of an event in the exposed group compared to that in the unexposed group. Odds give a sense of the risk (probability of an outcome among the total population able to experience the said outcome). The general logistic regression model in which there are multiple exposure variables is represented by the equation below:

Equation 3.6: Logistic regression

$$logit \quad 〖(y) = \beta\_0 〗 \; + \beta\_1 \, x\_1 + \beta\_2 \, x\_2 + \cdots \beta\_i \, x\_i$$

Where:

Logit y = Natural log of the odds ratio of outcome(y) given x

β_0 = Log of odds in the unexposed group

β = Log of odds associated with exposure variables x_1 to x_i

x = Exposure variables

Odds, unlike probabilities, are not constrained between 0 and 1 and can take values ranging from 0 to ∞. Therefore, the logistic regression model

can predict the probability of a binary outcome variable on a linear scale, simulating a linear regression.

In the logistic regression equation, parameters are estimated through a process known as maximum likelihood estimation, an iterative process through which the strongest linear combination of exposure variables increases the chances of detecting the observed outcome(Hosmer et al., 2013).

Logistic regression is governed by certain assumptions. One of these assumptions is the impendence of errors, meaning that all outcomes in the sample group must be independent of one another. Another assumption for continuous variables is a linear relationship with the logit of the outcome. The absence of multicollinearity and influential outliers are additional assumptions of linear regression models. In this thesis, assumptions were assessed using diagnostic plots and post-estimation analyses.

Following logistic regression modelling, calibration, or goodness of fit analyses to examine how well the models fit the data are conducted. Calibration analysis investigates the agreement between predicted and observed risk (Steyerberg et al., 2010). In this thesis, calibration was done using calibration plots. These plots display the predicted risk from the model on the x-axis and the outcome on the y-axis. A perfect prediction should fall on the 45-degree line. The calibration plot comprises an intercept or ‘calibration-in-the-large’, a measure of systematically too low or high predictions, and a calibration slope which should be as close to 1

as possible, with values smaller than 1 indicating an overfitted model (Steyerberg et al., 2010).

An accurate logistic regression model should discriminate between people with and without the outcome. In this thesis, this was assessed with the area under the Receiver Operating Characteristic (ROC) curve. Originating from signal detecting theory, this curve shows how the receiver detects signals in the presence of noise by plotting the probability of detecting the true signal (sensitivity) and false signal (1 – specificity) for a whole range of cut points. It evaluates the ability of the logistic regression model to assign higher outcome probabilities to subgroups who develop the outcome than to those who do not develop the outcome. Values range from 0.5 – 1.0. The area under the ROC curve value is interpreted using the following general rule of thumb. Values less than 0.5 indicate no discrimination, as assignment is more likely due to chance. Values greater than 0.5 but less than 0.7 indicate poor discrimination. ROC values greater than 0.7 but less than 0.8 are considered acceptable discrimination; greater than or equal to 0.8 and less than 0.9 depicts excellent discrimination and, values greater than or equal to 0.9 show outstanding discrimination(Hosmer et al., 2013).

For the purpose of this thesis, the relative importance of each variable to the regression model was of significance. Variable importance can be determined in several ways (Bi, 2012, Johnson and Lebreton, 2004). Wei et al provide a comprehensive overview of different variable important ascertainment methods (Wei et al., 2015). In this thesis, dominance analysis was employed. Dominance analysis is a technique that

examines the unique contribution of each variable to the prediction of the outcome variable by implementing pairwise comparisons of exposure variables in an index model across all subsets of exposure variables in the model (Azen and Traxel, 2009).

As mentioned earlier, logistic regression models estimate odds ratios (OR). An odds ratio of one occurs when the odds of disease in the exposed group is identical to that in the unexposed. ORs less than one mean the odds of the outcome is greater in the unexposed, and vice versa if the OR is greater than one. ORs are often interpreted as a risk ratio (relative risk - which is the ratio of the risk in the exposed group to that in the unexposed group) in the literature, but ORs only approximate risk ratios (or relative risk) when the outcome is rare. With common outcomes, the OR often overestimates the risk. That being said, estimating RRs is computationally intensive and, in the presence of multiple exposure variables challenging to interpret because the risk can only take values between zero and one, constraining the RR. Furthermore, RRs take on different values based on the choice of the outcome, i.e., the presence or absence of an event. In addition, RRs cannot be estimated in case-control studies. These issues are eliminated with the use of ORs. Considering the above, ORs reported for rare and common outcome involving multiple exposure variables is acceptable practice. In this thesis, ORs were presented.

3.3.5 Growth Models

The statistical methods described earlier require each observation in the sample population to be independent. This poses a unique challenge for

longitudinal data, with measurements at multiple time points because observations are not independent. For the purposes of this thesis, our interest lay in charting the developmental course of fatigue over time, as well as identifying heterogeneity in this development. Therefore, growth modelling was conducted. Growth modelling (also called growth curve modelling, trajectory modelling, and latent curve analysis) is a powerful and versatile statistical tool used to examine how attributes change over time(Grimm et al., 2016). There are several approaches to growth modelling, including multilevel modelling, and structural equation modelling, using latent variables. Growth models can also be used for continuous and non-continuous outcome variables, modelling linear and non-linear change. In this thesis, linear mixed effect modelling, and group-based trajectory modelling were used and therefore discussed further.

3.3.5.1 Linear mixed-effect modelling

This is also called multi-level linear or hierarchical growth curve analysis. Linear modelling is a simple and popular analytical technique and is often a common starting point in our attempt to statistically model a developmental process. Growth models can be estimated using a multilevel (hierarchical linear model) or structural equation modelling approach, with analogous results(Curran, 2003, Grimm et al., 2016, Sophia and Skrondal, 2012).

In the multilevel/hierarchical modelling, data are in various levels, for example, students within a school, or schools within a district, with each strata representing an analytical level. In growth curve modelling,

repeated measurements are nested within the individual, creating the first level of a multilevel model (level 1). The multilevel approach does not require a balanced design. Additionally, individuals may vary in their number of measurements (by design or attrition), making it a suitable analytical approach for the ERAN dataset. Grimm et al. provide an excellent overview of growth models and approaches for the applied researcher, including implementing these models using available statistical software(Grimm et al., 2016). A very brief description of linear growth modelling is provided below.

Basically, in its simplest form, the linear growth model is similar to the linear regression model described above. However, to adjust for interdependencies of individual measurements over time, a normal distribution is drawn around intercepts (the value of the outcome variable at the initial point in time) and the slopes (rate of change of outcome variable over time) of each individual in the population and these parameters are described by their means, variances and covariances(Twisk, 2013, Sophia and Skrondal, 2012). This process is represented in the equations below, where the random intercept model (no growth) represents the overall mean values of the outcome variable over time.

Equation 3.7: Random intercept model

$$Y\_it = \beta\_0i + \beta\_1\, X\_it + \varepsilon\_it$$

Where,

Y_{it} = observations for subject *i* at time *t*

β_{0i} = the random(over intercept,

X = covariate,

β_{1i} = the random regression coefficient for X, and

ε_{it} = the 'error' for subject *i* at time *t*.

The regression coefficients for covariates can also vary between subjects in longitudinal analysis, meaning that there is an interaction between covariates and the subject. This needs to be included in the model as well, and it is done in an analogous manner as adjusting for the within-subject interdependencies, i.e., a normal distribution is drawn around the different regression coefficients and variance estimated from this normal distribution. The variance is then added to the regression model. This is known as the random slope and is given in the equation below.

Equation 3.8: Random intercept random slope model

$$Y\_it = \beta\_0i + \beta\_1i\ X\_it + \varepsilon\_it$$

Where:

Y_{it} = observations for subject *i* at time *t*,

β_{0i} = s the random intercept,

X = the covariate,

β_{1i} = the random regression coefficient for X, and

ε_{it} = the 'error' for subject *i* at time *t*.

Growth curves can be non-linear, taking several shapes. These variations can be modelled using polynomial functions of time, for example, quadratic (addition of time squared to the model) or cubic

(addition of time cubed); or piecewise modelling, where time is split into segments, with breakpoints between segments also known as knots. Linear models for each segment are estimated. These linear models are also called splines(Sophia and Skrondal, 2012). In this thesis, linear, quadratic, and piecewise modelling were conducted.

Model selection was done by comparing nested models, and examining indices such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and -2 log-likelihood values. Lower values represent better model fit(Twisk, 2013).

The contribution of other variables to the variances and covariances of the intercepts and slope can be examined using hierarchical growth models. Variables that change over time, like pain score, are referred to as time-varying covariates, while those that do not change with time (e.g. sex) are known as time-invariant covariates(Grimm et al., 2016).

3.3.5.2 Group-based trajectory modelling (GBTM)

Linear growth models provide information on individual variability around a mean trajectory in a given population. However, other statistical methods are required to investigate trajectory heterogeneity within the population. One of such methods is called group-based trajectory modelling (GBTM). GBTM is part of a larger modelling framework called trajectory modelling. Simply put, it is a procedure that simplifies a heterogeneous population into more homogenous classes or clusters. This provides a framework to examine factors that determine the distinctiveness of these clusters or categories, so-called ‘deep

phenotyping' and facilitates the investigation of possible aetiological associations. Trajectory modelling is a relatively new technique in the field of epidemiology but has been used extensively in psychological research(Nagin, 2014, Nagin, 2005, Stevenson et al., 1996)

GBTM allows for repeated measures of a growth variable (in this thesis, fatigue) to be modelled as a function of a small number of latent growth factors and time-dependent latent errors. These latent growth factors are assumed to be independent and normally distributed and allow for the estimation of a mean trajectory and individuals' deviation from this trajectory. In addition, model parameters are allowed to vary across a specified number of latent groups (Gilthorpe et al., 2014, Lennon et al., 2018).

Practically, selecting the ideal number of latent classes is through an iterative process whereby different numbers of latent classes are specified each time and assessed against predetermined fit statistics. If a model with two or more latent classes fits the data better than a model with one latent class, it suggests that the population comprises subgroups, each with its own underlying "developmental process"(Gilthorpe et al., 2014). The choice of the best-fit model is also heavily dependent on the model parameterization matrices, i.e., which parameters will be estimated and which will be constrained (assigned a specific value by the researcher). Information required to make these decisions is heavily based on associated theory.

3.3.6 Missing Data

Missing data are a ubiquitous problem in epidemiological studies, particularly longitudinal studies. Missing data may be responsible for biased estimates and decrease study power by reducing the study sample size, invariably leading to wrong conclusions(Kang, 2013).

In longitudinal studies, missing data could be described in terms of pattern, whereby participants may miss a particular assessment during one follow-up visit (item missingness), an entire study assessment, and return to the study at a later timepoint (intermittent missingness), or never return to the study(drop-out). Item missingness, intermittent missingness and drop-outs usually co-occur in practice (Twisk, 2013).

Missing data mechanisms are often described in statistical literature according to assumptions governing the mechanisms of missing data(Rubin, 1976). According to Rubin, missing data could be missing completely at random (MCAR), where the probability that missing values of a variable are independent of the value of any other observed variable. This indicates that observed values are a random sample of the population with complete data and are therefore not systematically different from complete data. It is not possible to completely ascertain if data are MCAR. However, under this assumption, conventional methods of missing data handling yield unbiased estimates.

Missing data could be missing at random (MAR). Under this assumption, the probability of a missing variable is dependent on other observed variables in the dataset but not on the variable itself. Meaning that given

(conditioning on) the other variables, data are a random sample of the complete dataset.

Finally, Missing data could be missing not at random (MNAR), where missing data are dependent on unobserved data; therefore, missingness is related to events not measured by the researchers(Little and Rubin, 2014).

For correct interpretation of study results, given missing data, the researcher needs to decide if missing data are ignorable or non-ignorable. The above theoretical mechanisms of missing data facilitate this decision-making process. MCAR is usually classed as ignorable missing data. However, MAR and MNAR could be referred to as non-ignorable missing data and may require missing data handling techniques (Twisk, 2013).

There are several methods for handling missing data ranging from conventional methods such as listwise deletion, pairwise deletion and last observation carried forward (LOCF). These techniques are useful under the MCAR assumption but may produce biased estimates with MAR and MNAR (Schafer and Graham, 2002). More advanced techniques applicable in the context of longitudinal studies include multiple imputations (MI), full information maximum likelihood (FIML) or inverse probability weighting. In this thesis, missing data in the ERAN dataset were handled using multiple imputation using chained equations (MICE). This will be discussed further here.

MICE is implemented by selecting random values from a missing data distribution model and conditioning on the observed data using a defined model. This procedure is done several times, each time generating a separate complete dataset. These results are then combined into a single set of parameter estimates, standard errors, and test statistics. MI can be implemented using MNAR assumption, but typically, data are assumed to be MAR. One advantage of MICE is the ability to handle different types of variables, including continuous, binary, ordered and unordered categorical variables, because each variable is imputed using its own model. However, evidence of the utility of MICE is largely empirical(White et al., 2011). In this thesis, estimates obtained using MICE, analysis of complete case data and sensitivity analyses were conducted to facilitate robust conclusions.

3.4 Other methods

3.4.1 Non-response bias in population surveys

Surveys are a quick and inexpensive way to reach large segments of a population of interest. However, low response rates are a widely recognised problem of epidemiological surveys especially postal surveys such as KPIC (Groves et al., 1992). Non-response bias is described as "the mistake one expects to make in estimating a population characteristic based on a sample survey data in which, due to non-response, certain types of survey respondents are underrepresented"(Cheung et al., 2017). For example, if only young adults respond to a survey about the effect of national lock downs on mental

health, it would not be possible to generalise findings of this survey to the total population. Reasons for non-response are often very complex, cannot always be quantified, and are therefore hypothesised. Interestingly, higher response rates do not necessarily predict non-response bias. Previous studies have shown that females, older individuals, and people with higher education are more likely to participate in postal surveys. In such cases, non-response bias is associated with the general characteristics of the study and could be mitigated by calculating weighted proportions(Groves and Peytcheva, 2008, Tripepi et al., 2010a).

3.4.1.1 Generalising findings to the wider population/addressing non-response bias.

I was interested in fatigue prevalence in the general population therefore, it was important that findings form the analyses of KPIC data could be generalised to the general population of interest. The GP register is a good representation of the target population. However, for the purposes of this thesis, access to this register was not practicable. Therefore, in this sample, the UK population in the East Midlands was used as the target population to assess the representativeness of data, and the KPIC population back weighed to account for non-response bias.

Consequently, in this thesis, in addition to other methods described above, I weighted the KPIC sample to reflect the UK population in the East Midlands on some key variables using a procedure called Iterative proportional fitting (IPF) or raking.

Raking is a statistical method involving the creation of sampling weights in a given population so that it more accurately represents the wider population of interest. In raking, a set of variables in which the population distribution is known is selected, and weights assigned to each study participant iteratively until the sample distribution of our index population aligns with the wider population of interest. Raking is simple to implement and only requires knowledge of the marginal proportions of the wider population of interest, which in the case of KPIC data were proportions of key variables in the UK east midlands population. The raking procedure implemented in this thesis follows the recommendation described by DeBell and Krosnick (DeBell and Krosnick, 2009) and was implemented by the r statistical software package 'Anesrake' (Pasek and Pasek, 2018).

3.4.2 Patient and public involvement in research

Public and Patient involvement (PPI) in research is defined by INVOLVE as "*research being carried 'with' or 'by' members of the public rather than 'to', 'about', or 'for' them(*National Institute of Healthcare Research (NIHR), 2018*)*. INVOLVE clearly distinguishes between patient and public **participation**, which relates to members of the public taking part in a research study, and **Engagement**, involving the dissemination of information and knowledge about research by members of the public.

Where possible, PPI in this thesis was conceptualised using the framework described by Shippee and colleagues (Shippee et al., 2015). This framework utilizes a dynamic three-dimensional approach, which

involves patient and service user initiation – entry of lay people into the research process, building reciprocal relationships, where lay people involved in the study are viewed as equal partners in the research process, co-learning, where both researchers and lay people are educated on the research process, and PPI involvement respectively, and reassessment and feedback; the dynamic component of the framework that deals with clarifying the roles and expectations of the researcher and layperson and modifying as appropriate to meet the study objectives.

At the start of this PhD studentship, a PPI event was organised by the department of Academic Rheumatology to discuss research ideas with members of the public with the specific aim of shaping the research questions and research approaches. The development of a primary research study in chapter 6 of this thesis facilitated the utilisation of the three-dimensional approach to PPI described above. Details can be found in chapter 6.

3.5 Covid 19 Impact

The Covid 19 pandemic precluded the inclusion of an additional dataset from the Jyvaskyla rheumatoid arthritis registry in Finland. Data retrieval of this registry data was hindered by the Covid-19 pandemic and various lock-downs in the UK and in Finland. The proposed study protocol, ethical approval and email correspondence with the clinical lead of the Jyvaskyla team can be located in appendix 11.1.3 – 11.1.4. Briefly, the data are based on the RA registry of the Jyväskylä Central Hospital. All new

patients with suspected RA in the Central Finland health care district are referred to the hospital for diagnosis and management. The registry contains retrospective demographic and clinical data of RA patients from the 1980s and prospective data from 1995(Sokka, 2016). In addition to almost four decades of RA data, the Jyväskylä registry data would have provided the opportunity to compare RA prevalence rates in two geographical settings (UK and Finland), model changes in RA classification and management (for example, treat to target strategy and the introduction of biologic therapy) and the relationship with fatigue, as well as attempting to model heterogeneity in the progression of fatigue over a more extended period.

3.6 Summary

This chapter has provided an overview of the datasets and statistical methods used in this thesis. Also discussed are the strengths and limitations of the datasets. Techniques introduced in this chapter are used throughout the thesis. Specific applications or extensions to these methods are discussed in the respective chapters of this thesis.

The following chapter presents the first of three studies examining fatigue in RA and in the general population.

4. A Cross-sectional study of fatigue in rheumatoid arthritis

4.1 Outline

This chapter begins with a brief review of the literature and rationale for the study, followed by a description of the study methods, including statistical analyses employed and results obtained. An interpretation of study findings in the context of existing knowledge and areas for further research concludes this chapter.

4.2 Introduction

Although a dominant symptom in many musculoskeletal diseases, there is little published material on the prevalence of fatigue in early rheumatoid arthritis (RA). Prevalence quoted in the literature range from 20 to 97%(Walter et al., 2018, Genty et al., 2017, Diniz et al., 2017, Druce et al., 2015a, Gron et al., 2014, Van Hoogmoed et al., 2010, Pollard et al., 2006). The average disease duration in these populations ranged from five to 18 years. A study in a population with a disease duration of fewer than 12 months described the prevalence of fatigue as 45%(Walter et al., 2018). This study was based on a moderately sized sample of about 250 participants in the Netherlands. This number may not be representative of the general RA population because the country of residence has been shown to affect the prevalence of fatigue (Hifinger et al., 2016).

Cross-sectional studies revealed a range of factors associated with fatigue in RA. Female sex, pain, mental health, sleep, and functional capacity emerged as significant predictors of fatigue. The contribution of

these factors to fatigue in RA is, however, inconsistent. Pollard et al and van Hoogmoed et al did not find an association between sex and fatigue(Pollard et al., 2006, Van Hoogmoed et al., 2010). In other cross-sectional studies(Pollard et al., 2006, Bergman et al., 2009), the relationship between fatigue and disease activity was not elicited. Katz et al., and van Hoogmoed et al. did not find an association between fatigue and functional ability measured with Health Assessment Questionnaire (HAQ)(Katz et al., 2016, Van Hoogmoed et al., 2010). Pain was consistently associated with fatigue in all studies that assessed the relationship. Several studies did not find an association between inflammation and fatigue(Pollard et al., 2006, Bergman et al., 2009, Davis et al., 2008, Van Hoogmoed et al., 2010). In these studies, the average disease duration ranged from 10 – 21 years. Factors associated with fatigue in early RA have not been explicitly examined in the literature, and it is not clear if these may differ between early and established RA.

4.3 Aim and objectives

4.3.1 Aim

This study aims to investigate the prevalence of fatigue and factors associated with fatigue in a population with early RA.

4.3.2 Objectives

- Describe the prevalence of fatigue in early RA.
- Identify factors associated with fatigue in early RA.

4.4 Methods

4.4.1 Study design/source of data

Secondary analysis of baseline data from the early rheumatoid arthritis network (ERAN) was used to estimate the prevalence of RA and investigate the factors associated with fatigue in early RA. ERAN study regimen was described in chapter three.

4.4.1.1 Sample size considerations

As secondary data analysis, the study population was based on available data, so sample size calculation was not done before the study. However, in a post hoc analysis, assuming multiple linear regression, with nine predictors, alpha value of 0.05, 80% power, and base and full model R^2 values of 0.45 and 0.48, respectively, this study would require 280 participants. This sample size calculation was done using the STATA® version 15 “powerreg” command(Stata, 2017).

4.4.2 Exposures and outcomes

4.4.2.1 Outcome variable

As mentioned in chapter three, fatigue in ERAN was measured by the vitality subscale of SF36 (Vitality), normed to the UK general population on a 0 – 100 scale, where higher values represent a better quality of life (i.e. less fatigue). Vitality is a validated measure of fatigue in RA with good measurement properties(Cella et al., 2005). However, there is no standardized cut-off point to delineate the presence or absence of fatigue. Therefore, to estimate the prevalence of fatigue from Vitality, a cut-off point was set to one standard deviation (SD) below the mean

values for the total ERAN population. In addition, one SD (10 units) of Vitality corresponds to the minimally important clinical change of the normed score(Van Hoogmoed et al., 2010, Bjorner et al., 2007). Severe fatigue was classified as two SDs below the mean vitality.

4.4.2.2 Exposure variables

A detailed description of exposure variables has been provided in chapter three. A brief recap is provided below.

Demographic characteristics, including self-reported age at onset of disease (in years), sex (male/female), ethnicity (white/non-white), body mass index (kg/m2), smoking status (never, current, and ex), were assessed. Inflammation was measured using the erythrocyte sedimentation rate (ESR) in mm/hr.

Disease activity was measured using the disease activity score in 28 joints (DAS28) and its components, tender (TJC) and swollen joint counts (SJC) in 28 joints, each scaled 0 – 28, patient's global assessment of disease activity (PGA) was measured on a 0 – 100mm visual analogue scale and ESR. Seropositivity was classified according to the presence of rheumatoid factor and/or positive or borderline anti-cyclic citrullinated peptide antibodies (CCP). The presence or absence of nodules (y/n) on clinical examination. Comorbidities were assessed using the rheumatic disease comorbidity index (RDCI) on a scale of 0 – 10.

Pain and mental health were measured using the SF36 bodily pain (pain) and SF36 mental health (mental health) subscales, respectively. Like Vitality, these subscales were normalised to the UK population on a 0 –

100 scale with higher values representing a better quality of life, i.e., less pain and better psychological health.

Disability/functional limitation was measured using the health assessment questionnaire (HAQ) disability index (HAQ-DI), scored on a scale of 0 – 3, with higher scores indicating greater disability.

Medication use was represented by a binary (y/n) variable, with the affirmative comprising participants prescribed disease-modifying anti-rheumatic drugs (DMARDs). Disease duration was measured by the self-reported time since the onset of symptoms in months.

4.4.3 Statistical Analysis

The study population was described with means, standard deviations (SD), and median and inter-quartile ranges (IQR) for normally and non-normally distributed continuous variables, respectively. Normality was assessed graphically with histograms. Categorical variables were described using frequencies and proportions.

The crude prevalence of fatigue in early RA was calculated by dividing the number of people who reported fatigue (based on the definition above) by the total number of people in the ERAN baseline population with valid Vitality scores. The same approach was used to calculate the prevalence of severe fatigue in ERAN. Fatigue prevalence was standardized by age in 5-year bands and sex to the European standard population (ESP) 2013 to allow for comparison among different geographical settings. Sex standardization was conducted to account for the higher prevalence of fatigue in females in published studies.

A description of standardization, including the mathematical formula, can be found in chapter three. 95% confidence intervals were derived based on the assumption of a gamma distribution according to the method described by Fay and Feuer. This method explains the generation of confidence intervals for direct standardization. Fay and Feuer described gamma distribution-generated confidence intervals which are more robust than the standard normal distribution, especially where there are a small number of events (Fay and Feuer, 1997, Morris et al., 2018). ESP 2013 population values are provided in the appendix table 11.1. Standardized prevalence was estimated using the 'Epitools' package for the R statistical software(Aragon, 2020).

The association between baseline covariates and fatigue was examined using scatter plots and Spearman correlation coefficients (r). The resulting p values were adjusted using Bonferroni correction. Bivariate and multivariable linear regression with robust standard errors presenting standardized regression coefficients were used to examine the magnitude and direction of the association between fatigue and measured covariates in the ERAN dataset. Post estimation, linear regression diagnostics were examined graphically, and multicollinearity was investigated using variance inflation factor (VIF)(O'brien, 2007, Katz, 2011).

Sensitivity analysis was performed using a sub-sample of ERAN with a self-reported disease duration of six months or less.

Statistical analysis was conducted in STATA® versions 16 and 17(Stata, 2017, Stata, 2019) and R studio (R Core Team, 2021). All estimates were presented with 95% confidence intervals (CIs), and unless otherwise stated, significance levels were set to 0.05.

4.5 Results

The ERAN sample comprised 1236 participants, of whom 992 had Vitality scores recorded at baseline. The sample was predominantly female (68%) of white ethnicity (97%), with a mean age of 57 years (SD:14) and a mean Vitality score of 42 (SD:11). Approximately two-thirds of the sample met four or more 1987 ACR criteria for the classification of rheumatoid arthritis at baseline (67%). 61% of participants were seropositive. Mean pain and mental health scores at baseline were 34 (SD:11) and 46 (SD:11), respectively, and the median disease duration was 12 months (IQR:7 to 18). The mean baseline DAS28ESR was 4.7 (SD:1.6). Please note that higher values of pain, Vitality and mental health depict a better quality of life (less pain, less fatigue and better mental health, respectively). Table 4.1 describes the characteristics of the ERAN study population at baseline.

Table 4.1: Baseline characteristics of study population.

Characteristics, unit	**Values**	**n**
Vitality	41.82 (11.14)	992
Demographics		
Age, yrs	57.01 (14.03)	1,234
Female	839 (67.88) *	1,236
White ethnicity	1196 (97.00) *	1,233
BMI, Kg/M^2	26.84 (16.80 - 29.76)~	1,120
Never Smoker	477 (40.02) *	1,192
Ex-Smoker	404 (33.89) *	1,192
Current Smoker	311 (26.09) *	1,192
Disease Activity/Inflammation		
DAS28ESR	4.68 (1.56)	1,048
TJC	5.00 (2 to 11)~	1,230
SJC	4.00 (1 to 9)~	1,231
Seropositive	655 (61.04) *	1,073
Erosions Present	331(29.4) *	1,126
Nodule present	100 (9.52) *	1,050
Hb, mg/dl	13.10 (1.44)	1,204
4 or more 1987 ACR criteria met	832 (67.31) *	1,236
ESR, Mm/hr	24.00 (12 to 41)~	1,053
PROMS		
Pain	33.77 (10.63)	996
Mental Health	46.34 (11.15)	991
HAQ	1.08 (0.76)	1,200
PGA	43.50 (25.6)	1,231
Medication/Comorbidity		
Medication use	935 (78.31) *	1,194
RDCI	1.00 (0 to 2)~	1,195
Others		
Duration, Months	12.00 (7 to 18)~	1,236

Values are mean (SD) unless otherwise stated. Vitality - SF36 vitality, Age - age at onset, HAQ - health assessment questionnaire, PGA- patient's global assessment of disease activity, Hb - haemoglobin concentration, DAS28ESR - disease activity score (ESR), Pain - SF36 bodily pain, Mental Health - SF36 mental health, RDCI - rheumatic disease comorbidity index, Duration - disease duration in months, TJC - tender joint count, SJC - swollen joint count, ESR - erythrocyte sedimentation rate, BMI - body mass index, PROMs - patient-reported outcome measures, IQR - interquartile range. (* = n (**%**), ~=median(IQR))

4.5.1 Prevalence of fatigue at Baseline

78% (CI:71% to 86%) of the study population reported lower Vitality scores than the mean UK general population. The crude prevalence of fatigue at baseline was 45%, and the age and sex standardized prevalence of fatigue and severe fatigue at baseline were 44% (CI: 39% to 50%) and 19% (CI: 15% to 23%), respectively. Comparable prevalence were seen in people with a disease duration of six months or less (Table 4.2).

Table 4.2: Crude and age and sex standardized prevalence of fatigue in the ERAN study population at baseline

Description of population	Definition of fatigue	Crude Prevalence (%)	Standardised Prevalence (CI)%
Total ERAN population	Below Normed UK Mean	78	78 (71 to 86)
	Fatigue	45	44 (39 to 50)
	Severe Fatigue	20	19 (15 to 23)
Disease Duration of ≤ 6mths	Below Normed UK Mean	78	76 (62 to 92)
	Fatigue	51	51 (40 to 66)
	Severe Fatigue	19	19 (12 to 29)

Fatigue – SF36 Vitality ≤ 1 SD below UK population mean, Severe Fatigue - SF36 Vitality ≤ 2 SD below UK population mean, CI – 95% confidence interval

4.5.2 Baseline characteristics associated with fatigue.

Pain and mental health scores showed significant moderate positive correlations with Vitality scores at baseline ($r = 0.54$ & 0.61 respectively, each $p \leq 0.05$) i.e., a higher Vitality score was associated with higher SF36 bodily pain and mental health scores. Given that Vitality, SF36 pain and mental health were scored such that higher scores represented a better QoL; this result translates to a higher fatigue level being significantly

associated with worse pain and mental health. Conversely, functional ability measured using HAQ was negatively correlated with Vitality ($r = -0.50$, $p \leq 0.05$), translating to a direct relationship between functional ability and fatigue. DAS28ESR and its component indices, except PGA, were each weakly correlated with fatigue (Table 4.3).

Noteworthy from table 4.3 is the correlation between PGA and pain; worse pain (lower SF36 bodily pain) was associated with higher patient assessment of disease activity. Interestingly, disease duration was not significantly associated with fatigue, pain, disease activity or mental health. A correlation matrix scatterplot is presented in appendix figure 11.1.

In bivariate linear regression analysis (Table 4.4), characteristics such as sex, BMI, disease-related factors such as DAS28ESR and its component indices, and haemoglobin concentration were significantly associated with Vitality at baseline. Higher disease activity and inflammation scores were significantly associated with lower Vitality (more fatigue). However, seropositivity, presence of erosions and RF were not associated with Vitality. PROMS – mental health, pain, patient's global assessment of disease activity and functional capacity measured using HAQ were significantly associated with Vitality, with values representing a better QoL associated with higher Vitality scores. Disease duration was not significantly associated with Vitality.

Linear regression coefficients presented in the tables below (and in linear regression tables throughout this thesis) represent the change in Vitality

levels per unit change in respective characteristics. The unit column in each table shows the higher and lower limits of each characteristic and is not representative of the range of the regression coefficient presented.

Table 4.3:Pairwise correlation among baseline characteristics of ERAN population

	Values													
	Vitality	HAQ	PGA	ESR	Das28ESR	Pain	Mental health	Duration	BMI	Age	TJC	SJC	RDCI	Hb
Vitality	1													
HAQ	-0.5001*	1												
PGA	-0.4397*	0.5367*	1											
ESR	-0.1896*	0.3229*	0.2869*	1										
DASESR	-0.3807*	0.5786*	0.6679*	0.6064*	1									
Pain	0.5368*	-0.6217*	-0.6076*	-0.2633*	-0.5609*	1								
Mental health	0.6142*	-0.4081*	-0.3993*	-0.1576*	-0.2986*	0.4882*	1							
Duration	-0.003	0.0161	0.0209	-0.0026	0.0063	0.0328	0.0027	1						
BMI	-0.1125*	0.1399*	0.0568	-0.0254	0.0828*	-0.0649*	-0.0422	0.0056	1					
Age	0.0592	0.0932*	-0.0372	0.1764*	0.0624*	0.0289	0.1123*	-0.0915*	-0.0331	1				
TJC	-0.3462*	0.4766*	0.4562*	0.2220*	0.7872*	-0.4547*	-0.2560*	0.0263	0.0959*	-0.0272	1			
SJC	-0.1761*	0.3450*	0.3644*	0.3045*	0.6814*	-0.3077*	-0.1191*	-0.0534	0.0521	0.1101*	0.5662*	1		
RDCI	-0.1039*	0.1862*	0.0569*	0.0813*	0.0844*	-0.1219*	-0.0648*	-0.05	0.1244*	0.3670*	0.0556	0.0569*	1	
Hb	0.0963*	-0.1860*	-0.0878*	-0.4963*	-0.3083*	0.0719*	0.0721*	0.0084	0.1334*	-0.0726*	-0.1099*	-0.1675*	0.0165	1

Values are spearman's rho(r), * p≤ 0.05

Vitality - SF36 Vitality, Age - Age, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Hb - haemoglobin Concentration, DAS28ESR - Disease Activity Score (ESR), SF36BP - Normed SF36 Bodily Pain, SF36MH - Normed SF36Mental health, RDCI - Rheumatic Disease Comorbidity Index, Duration - Disease Duration in Months, TJC - Tender Joint Count, SJC - Swollen Joint Count, ESR - Erythrocyte Sedimentation rate, BMI - Body Mass Index.

Table 4.4: Characteristics associated with Vitality at baseline in the bivariable linear regression analysis.

Characteristic	Unit	Values (CI)
Demographics		
Age	years	0.04 (-0.01 to 0.09)
Female		-4.35 (-5.81 to -2.90) *
White Ethnicity		-3.22 (-7.40 to 0.97)
BMI	Kg/m^2	-0.24 (-0.37 to - 0.10) *
Never Smoker		Ref
Ex-Smoker		0.23 (-1.40 to 1.86)
Current-Smoker		-0.74 (-2.52 to 1.04)
Disease Activity/Inflammation		
DAS28ESR	0 - 9.4	-2.78 (-3.23 to -2.33) *
TJC	0 - 28	-0.56 (-0.66 to -0.48) *
SJC	0 - 28	-0.36 (-0.49 to -0.23) *
ESR	mm/hr	-0.09 (-0.12 to -0.06) *
Seropositive		-0.78 (-2.33 to 0.77)
Erosions		0.06 (-1.53 to 1.66)
Nodules		-1.10 (-3.58 to 1.38)
Hb	mg/dl	0.76 (0.26 to 1.26) *
PROMs		
SF36BP	0 - 100	0.56 (0.51 to 0.62) *
SF36MH	0 - 100	0.61 (0.56 to 0.66) *
HAQ	0 - 3	-7.39 (-8.19 to -6.59)*
PGA	0 - 100	-0.19 (-0.22 to -0.17) *
Comorbidities/Medication		
RDCI	0 - 9	-0.90 (-1.43 to -0.36) *
Medication use		1.49 (-0.24 to 3.22)
Others		
Duration	months	-0.003 (-0.06 to 0.06)

*$p \leq 0.05$. Values are linear regression coefficients.
SF36VT - SF36 Vitality, Age - Age at onset, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Hb - Haemoglobin Concentration, DAS28ESR - Disease Activity Score (ESR), SF36BP - SF36 Bodily Pain, SF36MH - SF36Mental health, RDCI - Rheumatic Disease Comorbidity Index, Duration - Disease Duration in Months, TJC - Tender Joint Count, SJC - Swollen Joint Count, ESR - Erythrocyte Sedimentation rate, BMI - Body Mass Index, PROMS – Patient Reported Outcome Measures, CI – 95% Confidence interval.

Table 4.5 shows the results of the multiple linear regression analysis investigating characteristics associated with fatigue at baseline. Sex, pain, mental health, HAQ, and tender joint count collectively explained about half of the variability of Vitality (R^2= 0.49). Females had on average lower Vitality scores than males. Worse pain, mental health, more tender joints, and worse functional ability were all associated with lower Vitality scores. Mental health scores showed the greatest impact on Vitality score (Standardized ß =0.42) compared with other characteristics in the model.

Table 4.5: Characteristics associated with Vitality at baseline in the multiple linear regression analysis.

Multivariable Linear Regression (n=973, R^2 = 0.49)			
Characteristic	Unit	β (CI)	Std β
Age	years	0.01 (-0.03 to 0.05)	0.01
Sex	female	-2.00 (-3.11 to -0.86) *	-0.08
TJC	0 - 28	-0.10 (-0.19 to -0.02) *	-0.06
Pain	0 - 100	0.21 (0.15 to 0.28) *	0.24
Mental health	0 - 100	0.42 (0.36 to 0.47) *	0.42
HAQ	0 - 3	-2.43 (-3.37 to -1.50) *	-0.17

*p ≤ 0.05, β – linear regression coefficient, Std – Standardised
Age - Age at onset, HAQ - Health Assessment Questionnaire, Pain - SF36 Bodily Pain, Mental Health - SF36Mental health, TJC - Tender Joint Count

Post-estimation linear regression diagnostics were acceptable (residual plots are shown in the appendix figure 11.3 & 11.4. The mean VIF was 1.14, meaning that multicollinearity was not elicited (Appendix table 11.2).

4.5.2.1 Sensitivity Analysis

Two hundred eighty-eight participants, representing 23% of the total ERAN population, had a disease duration of six months or less. Table 4.6

shows the demographic characteristics of this subgroup. Participants with a disease duration of less than six months did not differ from the rest of the ERAN population in terms of Vitality levels, ethnicity, BMI, and mental health scores ($p \geq 0.05$), as shown in table 6.7. This sub-group was older at disease onset, had higher inflammatory and disease activity scores, and slightly worse PROMS.

Table 4.6: Characteristics of study participants with disease duration ≤ 6 months

n=288		
Characteristics, unit	**Values**	**n**
Vitality	41.19 (10.68)	227
Demographics		
Age, yrs**.**	60.43 (14.73)	287
Female	180 (62.50) *	288
White ethnicity	283 (98.26) *	288
BMI, Kg/M^2	26.66 (23.87 - 30.03)~	264
Never Smoker	118 (41.99) *	118
Ex-Smoker	96 (34.16) *	96
Current Smoker	67 (23.84) *	67
Disease Activity/Inflammation		
DAS28ESR	5.00 (1.48)	251
TJC	6 (2 -12)~	288
SJC	6 (2- 10)~	288
Seropositive	146 (54.68) *	267
Erosions Present	59(22.10) *	267
Nodule present	26 (10.36) *	251
Hb, mg/dl	12.91 (1.37)	284
4 or more 1987 ACR criteria met	196 (68.06) *	288
ESR, Mm/hr	33.92 (24.51)~	251
PROMS		
Pain	31.25 (10.39)	228
Mental	45.07(10.72)	227
HAQ	1.19 (0.75)	283
PGA	48.10 (25.23)	288
Medication/Comorbidity		
Medication use	200 (73.80) *	271
RDCI (Median, IQR)	1.00 (0 to 2)~	271
Others		
Duration, Months	5 (4 to 6)~	288

Values are mean (SD) unless otherwise stated. (* = n (%), ~=median (IQR))
Vitality - SF36 Vitality, Age - Age at onset, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Hb - haemoglobin Concentration, DAS28ESR - Disease Activity Score (ESR), SF36BP - SF36 Bodily Pain, SF36MH - SF36Mental health, RDCI - Rheumatic Disease Comorbidity Index, Duration - Disease Duration in Months, TJC - Tender Joint Count, SJC - Swollen Joint Count, ESR - Erythrocyte Sedimentation rate, BMI - Body Mass Index, PROMs -Patient-Reported Outcome Measures, IQR - Inter Quartile Range.

Table 4.7:ERAN population with a disease duration greater than six months compared with the population with a disease duration of six months or less.

Characteristics, unit	≤6 months (n-288)	>6 months (n-948)	P value
Vitality	41.19 (10.68)	42.01(11.27)	0.3266
Demographics			
Age, yrs**.**	60.43 (14.73)	56.13 (13.68)	<0.0001
Female	180 (62.50) *	659 (69.51) *	0.026
White ethnicity	283 (98.26) *	913 (96.61) *	0.151
BMI, Kg/M^2 (Median, IQR)	26.66 (23.87 to 30.03)	26.85 (23.95 to 30.82)	0.1527
Never Smoker	118 (41.99) *	359 (39.41) *	
Ex-Smoker	99 (34.16) *	308 (33.81) *	
Current Smoker	67 (23.84) *	244 (26.78) *	0.3122~
Disease Activity/Inflammation			
DAS28ESR	5.00 (1.48)	4.58 (1.57)	0.0002
TJC (Median, IQR)	6 (2 to 12)	5 (1 to 11)	0.0428
SJC (Median, IQR)	6 (2 to 10)	4 (1 to 8)	<0.0001
Seropositive	146 (54.68) *	509 (63.15) *	0.014
Erosions Present	59 (22.10) *	272 (31.66) *	0.003
Nodule present	26 (10.36) *	74 (9.26) *	0.606
Hb, mg/dl	12.91 (1.39)	13.16 (1.45)	0.0117
4 or more 1987 ACR criteria met	196 (68.06) *	636 (67.09) *	0.759
ESR, Mm/hr (Median, IQR)	33.92 (24.51)	29.07 (23.74)	0.0052
PROMS			
Pain	31.25 (10.39)	35.52 (10.59)	<0.0001
Mental health	45.07 (10.71)	46.71 (11.25)	0.0504
HAQ	1.19 (0.75)	1.05 (0.76)	0.0082
PGA	48.10 (25.22)	42.09 (25.58)	0.0005
Medication/Comorbidity			
Medication use	200 (73.80) *	735 (79.63) *	0.041
RDCI (Median, IQR)	1 (0 to 2)	1 (0 to 2)	0.0255
Others			
Duration, Months (Median, IQR)	5 (4 to 6)	14.5 (10 to 26)	

P values obtained from Ttest, Man U or Chi Squared tests, ~p value for trend**. (*** = n,%)
Values are mean (SD) unless otherwise stated. SF36VT - SF36 Vitality, Age - Age at onset, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Hb - haemoglobin Concentration, DAS28ESR - Disease Activity Score (ESR), Pain - SF36 Bodily Pain, mental health - SF36Mental health, RDCI - Rheumatic Disease Comorbidity Index, Duration - Disease Duration in Months, TJC - Tender Joint Count, SJC - Swollen Joint Count, ESR - Erythrocyte Sedimentation rate, BMI - Body Mass Index, PROMs -Patient Reported Outcome Measure, IQR - Inter Quartile Range.

Table 4.8 below shows the characteristics associated with fatigue at baseline in the sensitivity analysis. Similar to the main analysis, female gender, ESR, PROMS and comorbidities were associated with Vitality in the bivariate analysis. In multivariable analysis (Table 4.9), worse pain, functional capacity and mental health were found to be associated with fatigue and explained about half of fatigue variability (R^2= 0.47).

Table 4.8: Bivariate linear regression in people with disease duration of ≤ 6 months

Characteristic	Unit	Values (CI)
Demographics		
Age	years	0.16 (-0.08 to 1.11)
Female		-3.46 (-6.30 to -0.61) *
White Ethnicity		-9.58 (-20.14 to 0.99)
BMI	Kg/m^2	-0.21 (-0.50 to 0.08)
Never Smoker		Ref
Ex-Smoker		1.81 (-1.40 to 5.02)
Current-Smoker		2.64 (-1.00 to 6.27)
Disease Activity/Inflammation		
DAS28ESR	0 - 9.4	-2.31 (-3.28 to -1.34) *
TJC	0 - 28	-0.37 (-0.56 to -0.17) *
SJC	0 - 28	-0.16 (-0.41 to 0.08)
ESR	mm/hr	-0.09 (-0.14 to -0.02) *
Seropositive		0.02 (-2.88 to 2.92)
Erosions		0.06 (-1.53 to 1.66)
Nodules		3.61 (-1.23 to 8.44)
Hb	mg/dl	0.55 (-0.49 to 1.58)
PROMS		
Pain	0 - 100	0.51 (0.39 to 0.63) *
Mental health	0 - 100	0.64 (0.54 to 0.74) *
PGA	0 - 100	-0.19 (-0.24 to -0.14) *
Hb	mg/dl	0.76 (0.26 to 1.26) *
HAQ	0 - 3	-6.11 (-7.84 to -4.40) *
Comorbidities/Medication		
RDCI	0 – 9	-0.81 (-1.84 to -0.21) *
Medication use		3.03 (-0.17 to 6.23)
Others		
Duration	months	-0.40 (-0.62 to 1.44)

*p ≤ 0.05 Values are Linear regression coefficients, CI - 95% Confidence Interval
SF36VT - SF36 Vitality, Age - Age at onset, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Hb - haemoglobin Concentration, DAS28ESR - Disease Activity Score (ESR), Pain - SF36 Bodily Pain, Mental health - SF36Mental health, RDCI - Rheumatic Disease Comorbidity Index, Duration - Disease Duration in Months, TJC - Tender Joint Count, SJC - Swollen Joint Count, ESR - Erythrocyte Sedimentation rate, BMI - Body Mass Index, PROMS – Patient Reported Outcome Measures

Table 4.9:Multivariable linear regression model in people with disease duration ≤ 6 months.

Multivariable Linear Regression (n=225, R^2 = 0.47)			
Characteristic	Unit	β (CI)	Std β
Age	years	0.002 (-0.07 to 0.08)	0.002
Sex	female	-1.85 (-4.13 to -0.42)*	-0.09
Pain	0 - 100	0.19 (0.04 to 0.35) *	0.19
Mental health	0 - 100	0.49 (0.38 to 0.60) *	0.50
HAQ	0 - 3	-2.02 (-3.94 to -1.00) *	-0.14

*p ≤ 0.05, β - Linear regression coefficients. Note that higher values of SF36 bodily pain and mental health represent a better quality of life, Vitality - SF36 Vitality, Age - Age at onset, HAQ - Health Assessment Questionnaire, Pain - SF36 Bodily Pain, Mental Health - SF36Mental health, CI – 95% Confidence interval.

4.5.2.2 Model Diagnostics

Graphical examination of residuals from multivariable linear regression models of both the total population and the subgroup revealed that linear regression assumptions were not violated. See appendix figure 11.5. Residuals from both models were normally distributed, appendix figure 11.6). Multicollinearity was examined in both multivariable models, revealing a mean VIF of 1.41 and 1.47, respectively, appendix table 11.3.

4.6 Discussion

4.6.1 Key findings and interpretations

This study investigated the prevalence of fatigue and the factors associated with fatigue using baseline data from a UK-based inception cohort of people with early RA. Key findings were that fatigue was prevalent in early RA and that female gender, pain, mental health, and functional capacity were significantly associated with fatigue in early RA. In addition, inflammation and other disease-associated factors were not significantly associated with fatigue in early RA. Another interesting

finding was that mental health was the most important determinant of fatigue in early RA.

The prevalence of fatigue in early RA is consistent with several published studies on people with established RA and different geographical settings (Diniz et al., 2017, Van Hoogmoed et al., 2010, Walter et al., 2018, Gron et al., 2014). Findings from this study corroborate other studies on the association between pain, mental health and functional capacity, and fatigue(Pollard et al., 2006). These associations were elicited in the total population and in the subgroup of people with disease activity of six months or less. It could therefore be inferred that disease duration is not an important determinant of fatigue in RA, and that fatigue is a problem even in early RA.

Consistent prevalences of fatigue in early RA (elicited in our study) and established RA (cited in the literature) (Genty et al., 2017, Gron et al., 2014, Van Hoogmoed et al., 2010), as well as the absence of a significant association between fatigue and measures of disease activity and inflammation, may signify a persistence of fatigue in people managed with intensive anti-inflammatory therapies. In contrast, short-term clinical trials showed a reduction in fatigue levels following biotherapies(Chauffier et al., 2012). This requires further investigation using longitudinal study designs.

In our study, mental health showed the highest contribution to fatigue variability. The association between mental health and fatigue is well documented in the literature (Williamson et al., 2005, Corfield et al., 2016,

Feldthusen et al., 2016b). Fatigue and depression are highly comorbid conditions in many inflammatory diseases. It has been theorized that conditions associated with unabated chronic inflammation, such as autoimmune diseases and certain cancers, lead to changes in neuronal function and, consequently, clinical depression in vulnerable people(Dantzer et al., 2008). Although clinically separate entities, it is essential to note that fatigue is a symptom in the diagnostic criteria of depression(Gruenberg et al., 2005).

Further supporting the results of this study are previous studies that have demonstrated a link between pain and fatigue in RA. RA pain is complex and heterogeneous and may vary in early or established RA (Mathias et al., 2021). Analogous to fatigue, RA pain has also been found to persist despite intensive treatment of inflammatory disease and is associated with poor sleep and low mood. In addition, one study evaluating pain and fatigue in RA found a synchronous rather than a causal relationship; therefore, the associations may be explained by a shared aetiology rather than a direct connection.

As previously highlighted, and in contrast to some published studies, I did not find an association between fatigue and inflammation or disease activity measures. This, perhaps, could be explained by the use of composite scores such as DAS28 as a measure of disease activity rather than individual components as was used in this study(Stebbings et al., Bergman et al., 2009). Composite scores include measures of inflammation and PROMs, which other factors outside inflammation could influence. This was recognised in people with fibromyalgia, who

had similar overall DAS28 scores to RA patients, even though statistically significant differences in component scores were elicited(Leeb et al., 2004).

Age of onset was not significantly associated with fatigue in our study. Previous research has shown an inconsistent relationship between fatigue and age (Belza et al., 1993, Olsen et al., 2016, Van Hoogmoed et al., 2010). The association between age and fatigue in some of these studies could have resulted from an indirect effect through other measures, such as frailty and comorbidities, rather than a direct effect of age. For example, in our dataset, frailty data were not collected; therefore, the contribution of frailty to fatigue could not be elicited. We did, however elicit a small mediational effect of comorbidities on the association between age and fatigue in our post hoc mediation analysis (Appendix 11.4.1).

4.6.2 Strengths and limitations

This study provided a cross-sectional examination of the prevalence of fatigue at baseline and factors associated with fatigue in a large population with early RA. Fatigue was measured using a validated instrument, and inflammatory and non-inflammatory characteristics were considered. Robust statistical analysis was conducted, which can allow for comparison among geographical settings. It is important, however, to interpret study findings in the context of certain caveats. In addition to the limitations of the ERAN dataset described in chapter 3, this is a cross-sectional observational study. As such, causal associations cannot be established using this study design. Another limitation of the study design

is bias caused by omitted confounders and/or residual confounding(Fewell et al., 2007). One of such omitted confounders is sleep. Sleep has been shown to be associated with pain, fatigue, and mental health(Whibley et al., 2019, Finan et al., 2013, Baglioni et al., 2016, Hammam et al., 2020). Sleep data was not collected in the ERAN cohort, so assessing the contribution of sleep to RA fatigue in this sample was not possible. The presence of sleep data could have explained more fatigue variability in the study population. Another important caveat is the lack of standardized cutoff values to delineate the presence or absence of fatigue. Clearly, not everyone experiences fatigue and estimated prevalence depend on the cut-off value applied. To mitigate bias that arbitrary cutoff values may introduce, I based the value on both the MCID and the UK population average scores.

4.7 Conclusion

Fatigue is prevalent in early RA and is associated with patient-reported outcomes such as mental health, pain and functional ability. Longitudinal studies are required to assess the progression of fatigue in RA and investigate the presence or absence of heterogeneity in the natural history of fatigue in RA.

4.8 Summary

In this chapter, I begin to investigate fatigue in early RA with a cross-sectional study. The prevalence and factors associated with fatigue were described. In the next chapter, I describe the course of fatigue in early

RA, factors associated with the course of fatigue and explore different courses of fatigue in the ERAN study population.

5. Longitudinal course of fatigue in RA

5.1 Outline

This chapter examines the longitudinal course of fatigue in early RA. Data from the early rheumatoid arthritis network (ERAN) was used to chart the progression of fatigue over time. Is there is a change in fatigue levels over time? And Investigate if this change (if present) is constant or a function of other characteristics such as disease duration, pain etc.

Trajectory heterogeneity was assessed to investigate if a single trajectory describes the entire population or if multiple groups of people exist with distinct trajectories. The chapter concludes with an explanation of study findings in the context of available knowledge and suggests areas for more research.

5.2 Introduction

Fatigue is a prevalent symptom in people with RA. (see previous chapters). Chapter 4 identified some patient-reported outcome measures (PROMs) as risk factors for fatigue in the RA population. Inflammation and disease activity were not significantly associated with fatigue. The progression of fatigue, however, is less well described. Available longitudinal studies focus on characteristics and conditions associated with persistent fatigue but do not explore alternative trajectory shapes; for example, a 'J-shaped' trajectory has been described for functional limitation early in the disease course(Norton et al., 2013). It is unclear if fatigue follows a similar trajectory or if certain factors modulate the course of fatigue.

Additionally, some available longitudinal studies examine drug effects on fatigue as a secondary outcome(Almeida et al., 2016) and report mean population effects that may mask individual subgroups based on fatigue trajectories. Population heterogeneity has been described for pain and functional limitation in early RA(Norton et al., 2013, McWilliams et al., 2019). Other longitudinal studies investigating fatigue trajectories often focus on a select population of people receiving biologic therapy, usually reserved for individuals who do not respond to conventional DMARDs (Provan et al., 2020, Bullock et al., 2018). Findings resulting from studies on these apriori selected populations may not be generalisable to the wider RA populations.

This study, therefore, seeks to investigate the longitudinal course of fatigue in a population of individuals with early RA. I hypothesise that individuals with RA make up a heterogeneous population based on fatigue trajectories, and the groupings could be identified early in the course of the disease. Identifying these heterogeneous groups may represent potential target populations for RA fatigue-related research and interventions.

5.3 Aims and Objectives

5.3.1 Aim

The study will investigate the course of fatigue in individuals with early RA.

5.3.2 Objectives

- To describe the longitudinal course of fatigue and identify predictors of change in fatigue in people with early RA
- To investigate fatigue trajectory heterogeneity
- To identify risk factors associated with fatigue trajectories (if any)

5.4 Methods

5.4.1 Study Design/Source of data

The study population used in this analysis consisted of a sub-population of the ERAN cohort (chapter 3), who recorded three or more SF36 vitality measurements.

5.4.2 Exposures and outcomes

In addition to ERAN exposure variables described in chapter 4, some covariates were modified as follows. Continuous variables with no meaningful value of "0", such as BMI, were mean-centred to ease interpretation(Enders and Tofighi, 2007).

Time in this study was assessed using measurement occasion. The first measurement occasion was defined as that at baseline, and the second was at three to six months. Subsequent measurements were conducted annually from baseline. Data were used up to the fifth measurement occasion (3 years from baseline).

5.4.3 Statistical Analysis

Continuous variables were described using means and standard deviations, and median and inter-quartile ranges for normally and non-

normally distributed covariates, respectively. Normality was assessed graphically with histograms. Categorical variables were described using frequencies and proportions. Scatter plots of growth patterns for fifty randomly selected individuals in the dataset were presented as well to give a broad graphical overview of the growth pattern.

The progression of fatigue over time was assessed using hierarchical growth curve modelling (GCM). This method was introduced earlier (chapter 3). GCMs model individual growth trajectories over time and describe how these trajectories vary both randomly or systematically as a function of time or other covariates(Curran et al., 2010). Growth curve models could be assessed using the multilevel (hierarchical growth curve) or structural equation modelling framework, as discussed previously. Both methods are empirically and analytically equivalent in studies of repeated observations (Curran, 2003).In this study, growth curve modelling was conducted using the multilevel modelling framework as follows:

First, the overall fatigue trajectory was investigated using the unconditional growth model, i.e., a growth model without covariates. A random intercept for participants was incorporated in this model to capture variability in starting values of fatigue, and random slopes for person and time were added to allow for variability in the rate of change for each study participant over time. The model with the best-fit statistics was selected for further analysis. I investigated a first-order autoregressive correlation, which assumes that Vitality measurements closer together were more correlated to measurements further apart. Fit

statistics used included the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). AIC provides an estimate of the information lost when a specific model is used to represent the data-generating process. It can be used to compare nested models, and the model with the least AIC value is usually chosen as the best fit because it represents a good balance between the goodness of fit of the model to the data and the model complexity(Cavanaugh and Neath, 2019, Profillidis and Botzoris, 2019). BIC is closely related to the AIC; however, it is favoured in larger samples and measures the efficiency of the model in terms of predicting the data. The BIC penalizes the complexity of the model and consequently tends to choose more parsimonious (models with fewer variables) models(Neath and Cavanaugh, 2012). Following the selection of the best fitted unconditional model, covariates were introduced into the model and eliminated by statistical significance (West et al., 2006). Interaction terms were included to assess the association between covariates and time. Only significant effects ($p \leq 0.05$) were included in the final model.

Missing data were assumed to be primarily due to attrition, as the proportion of missing data for each variable increased with each measurement occasion. Missing data were handled with multiple imputation using chained equations (MICE). Following examination of the patterns of missing data for each wave (chapter 3), data were assumed to be missing at random. MICE was implemented using the ***miceadds*** package of R statistical software, which allows for modelling the longitudinal structure of the data(Robitzsch et al., 2017). One hundred

imputed datasets were created. Convergence was assessed graphically using density plots, and ten burn-in iterations were selected. Imputed datasets were combined using Rubin's rules and compared to values obtained using list-wise deletion(Manly and Wells, 2015). Imputed values are presented here, and values from the listwise deletion analysis and convergence plots are presented in the appendix.

Fatigue trajectory heterogeneity was investigated using group-based trajectory modelling (GBTM), chapter 3. GBTM is one of a group of statistical models aimed at identifying distinct developmental progressions that may not be identified using conventional methods such as cluster analysis. Clusters of individuals who follow a similar developmental trajectory are selected by the probabilities of group membership(Nagin, 2014).

Identification of the number of groups was done through an iterative process recommended by Nagin (Nagin, 2005) and reported according to the guideline for reporting on latent trajectory studies (GRoLTs) checklist (van de Schoot et al., 2017), appendix table 11.4. One, two and three group models were tested in turn, following which linear and quadratic trajectories were examined. Starting values were obtained from the parameter estimates of the single trajectory model to avoid convergence issues. The most parsimonious model with the best fit statistics was selected. Fit statistics included the AIC and BIC, where lower values indicated better fit. Additionally, Entropy, an indication of how well groups are separated and how well individuals fit into their respective groups, was used. Entropy values closer to 1 on a range of 0 – 1 indicate better

precision(Nagin and Odgers, 2010). Other statistics examined in the final model included the proportion of individuals assigned to each group. A proportion ≤ 1% assigned to the smallest group indicates poor fit (Klijn et al., 2017). The average posterior probability (APP) is defined as the average of the posterior probability of group membership. The posterior probability is the probability of an individual's membership in an index group. A minimum of 0.7 is considered acceptable. The odds of correct classification, of which higher values indicate a better model fit and values less than five for all groups indicates good assignment precision was calculated (Nagin and Odgers, 2010). GBTM was conducted using the STATA® *"Traj"* plugin (Jones and Nagin, 2013). In accordance with the GRoLTS checklist, an alternative specification using growth mixture modelling (GMM) was conducted using the "*lcmm*" package in R version 4.0.5 (Proust-Lima et al., 2021).

Following group classification based on maximum posterior probability, each class was treated as a distinct entity. This means each participant in the dataset was assigned to the group with the largest posterior probability. Based on this classification, factors associated with group membership can be assessed using conventional statistical techniques, described as the classify-analyse strategy(Klijn et al., 2017). In this study, a logistic regression model was created with group membership as the binary outcome variable. Prior to modelling, the baseline characteristics of each group were described using simple descriptive statistics. Bivariate and multivariable logistic regression were employed. Following

model selection, discrimination was examined using the area under the receiver operating characteristic curve (AUROC).

5.4.3.1 Sensitivity Analysis

GBTM handles missing data through maximum likelihood estimation, which produces unbiased estimates under the MAR assumption(Klijn et al., 2017). However, sensitivity analysis was conducted to interrogate the validity of the results. This was conducted in two ways. Firstly, the analysis was restricted to people with a disease duration of six months or less, and secondly, to people who had five complete Vitality measurements.

P values were set at ≤ 0.05, and statistical analysis was conducted using STATA® 16 unless otherwise stated.

5.5 Results

5.5.1 Overview of the study population

Of 1236 ERAN participants at baseline, 729 participants had at least three Vitality measurements. Figure 5.1 shows a flow diagram of ERAN participants used in the analyses. 73.4% of the total ERAN population who had a Vitality recording at baseline (992) were included in the longitudinal analysis.

Descriptive characteristics of the population used in the longitudinal analysis at baseline are presented in table 5.1 below. The mean Vitality score was 42 (SD:11), and the median disease duration was 13 months (IQR: 7 to 22). The population was predominantly female (68%), with

over 90% of white ethnic origin. The mean DAS28ESR was 4.79 (SD:1.54), signifying moderate disease activity. Pain and mental health scores were 33(SD:11) and 47(SD:11), respectively, both lower than the general population mean value of 50. This population did not differ from the total baseline ERAN population in terms of measured covariates, table 5.2. The distribution of covariates over the measurement occasions is presented in the appendix table 11.5.

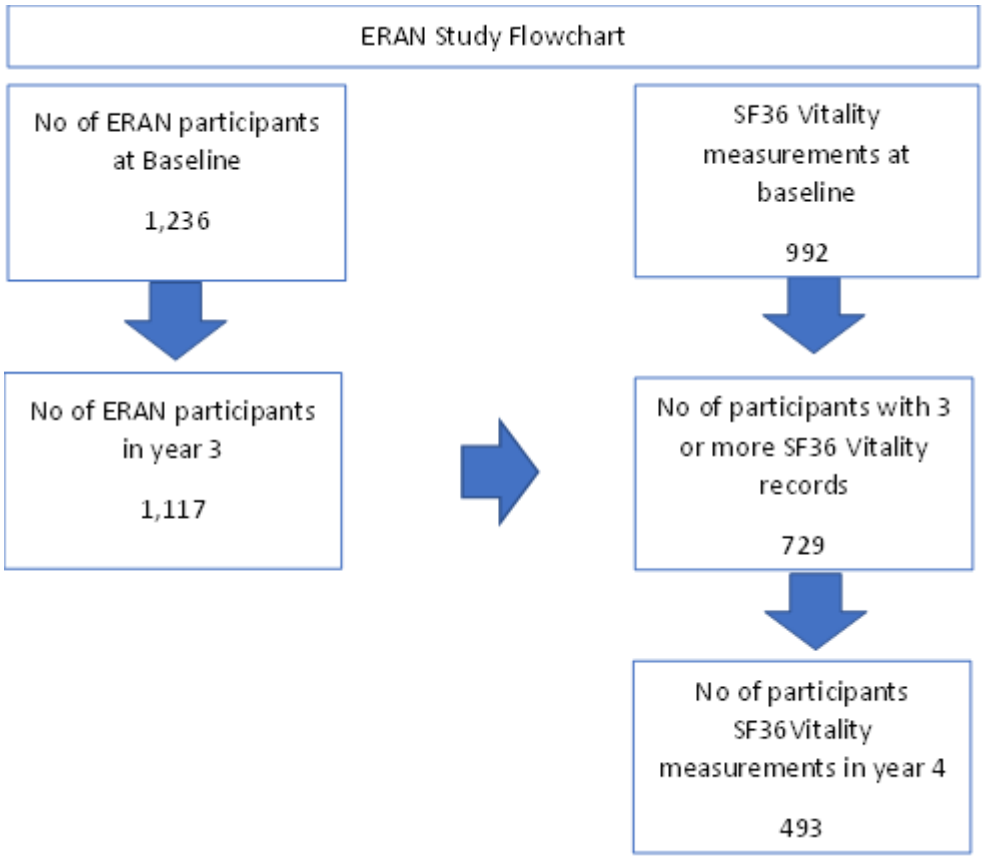


Figure 5.1: ERAN study participant flow chart

Table 5.1: Descriptive characteristics of the ERAN population with ≥ 3 Vitality measurements (n=729)

Characteristics, unit	**Three or more measurements**	**n**
SF36VT	42.13 (11.18)	664
Demographics		
Age, yrs.	56.55 (13.50)	729
Female	493 (67.63) *	729
White ethnicity	709 (97.26) *	725
BMI, Kg/M^2	27.05 (24.14 - 30.47)~	686
Never Smoker	280 (39.60) *	280
Ex-Smoker	240 (33.95) *	240
Current Smoker	187 (26.45) *	187
Disease Activity/Inflammation		
DAS28ESR	4.79 (1.54)	612
TJC	6 (2 to 12)~	726
SJC	5 (2 to 10) ~	726
Seropositive	389 (62.74) *	620
Erosions Present	213 (31.56) *	675
Nodule present	64 (9.70) *	660
Hb, mg/dl	13.11 (1.43)	706
ESR, Mm/hr	24 (12 to 42)	614
PROMS		
SF36BP	33.32 (10.63)	665
SF36MH	46.92 (11.17)	663
HAQ	1.11 (0.76)	717
PGA	43.85 (24.93)	727
Medication/Comorbidity		
Medication use	609 (83.54) *	729
RDCI	1 (0 to 2)~	729
Others		
Duration, Months	13 (7 to 22)~	729

Values are mean (SD) unless otherwise stated. SF36VT - SF36 Vitality, Age - Age at onset, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Hb - haemoglobin Concentration, DAS28ESR - Disease Activity Score (ESR), Pain - SF36 Bodily Pain, Mental health - SF36Mental health, RDCI - Rheumatic Disease Comorbidity Index, Duration - Disease Duration in Months, TJC - Tender Joint Count, SJC - Swollen Joint Count, ESR - Erythrocyte Sedimentation rate, BMI - Body Mass Index, PROMs -Patient Reported Outcome Measure, IQR - Inter Quartile Range. (* = n(%), ~=median(IQR))

The progression of Vitality over the study period is presented graphically in Figure 5.2 below for a random selection of 50 study participants. The data does not show a clear pattern of change over time; trajectories appear flat, with some variation between measurement occasions in some individuals.

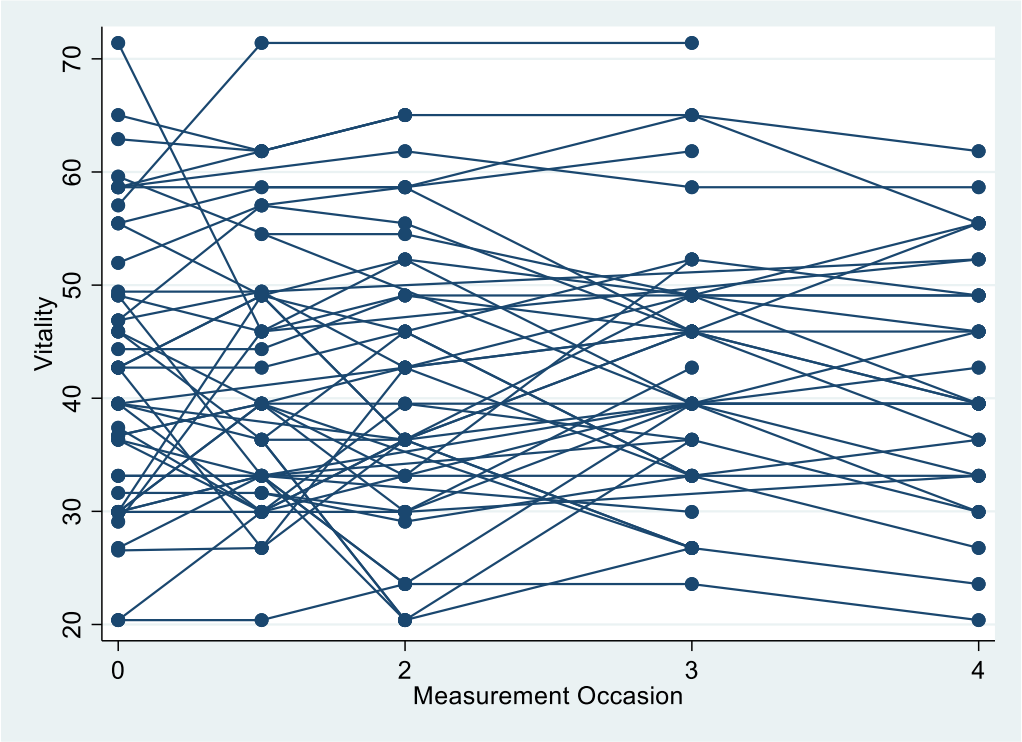


Figure 5.2: Spaghetti plot showing the progression of Vitality over time in 50 randomly selected ERAN participants.

Table 5.3 presents estimates of the progression of fatigue over time, assuming a constant change over time per individual, a random intercept, with varying rates of change per individual (linear growth), quadratic, and piecewise change over time, with knots at the first measurement occasion (six months after baseline), and second measurement occasion (1 year after baseline), and some fit statistics including AIC, BIC and -

2LL. The table shows that the random intercept model closely resembles the linear growth model in terms of estimates and fit statistics.

5.5.2 Progression of fatigue over time.

Table 5.2: Table showing the progression of Vitality over time, using a random intercept only, random intercept/random slope, quadratic, and linear spline with internal nodes at the first & second measurement occasion (corresponding to 6 months & 1 year after baseline respectively)

		Values				
		Random Intercept	Random Intercept/Random Slope	Quadratic growth	Spline1	Spline2
Model Parameters	Time	0.192* (0.06 to 0.32)	0.192* (0.05 to 0.33)	0.85* (0.37 - 1.32)	1.04* (0.25 to 1.83)	
	$Time^2$			-0.11* (-0.18 to -0.03)		
	S_1				-0.40 (-1.29 to 0.49)	1.12* (0.34 to 1.89)
	S_2					0.10 (-0.13 to 0.33)
	Constant	42.32* (41.60 to 43.04)	42.32* (41.58 to 43.06)	41.85* (41.04 to 42.66)	41.79* (40.97 to 42.60)	41.79* (40.97 to 42.60)
Fit Statistic	Log Likelihood	-13769.28	-13749.60	-13739.14	-13740.28	-13740.49
	Akaike Inf. Crit.	27546.57	27511.20	27498.27	27500.57	27500.98
	Bayesian Inf. Crit.	27571.38	27548.40	27560.28	27562.58	27562.99

*p<0.05; Values are SF36Vitality change per measurement occasion.
No Growth Random intercept only model, linear growth – linear growth model, quadratic growth – quadratic growth model, Spline 1 – Linear splines with a knot at second measurement occasion(s_1), Spline 2 – splines with a knot at first and second (s_1&s_2)measurement occasions.

In the random intercept model, the rate of change was 0.19 (CI:0.05 to 0.33), meaning that Vitality increased by 0.19 points per measurement occasion. The random intercept/random slope model showed a similar rate of change of 0.19 (CI:0.05 to 0.33) but with slightly better fit statistics (lower AIC and BIC). The quadratic and piecewise models had poorer fit statistics than the random intercept/random slope models. Values for non-imputed data showed similar results, appendix table 11.6. The first-order autoregressive estimate did not show a significant correlation, rho = 0.01 (CI:-0.1 to 0.12).

In the fully conditional model (model including covariates), table 5.3, Vitality was found to remain relatively stable through the period of study with little variability between measurement occasions, standardised ß = -0.02 (CI: -0.04 to -0.00). Random effects of the model above, table 5.3 showed that in 95% of study participants, Vitality values varied between 12.23 to 34.10 ($\tau 00$= 23.27) at baseline, with the rate of change over the study period varying between -1.21 and 0.90 ($\tau 11$ = 0.27). Additionally, there was a negative intercept slope covariance ($\rho 01$=-0.60), meaning that people with lower baseline Vitality levels at baseline tended to have a higher rate of change over time. 48% of the variability in Vitality was due to differences within study participants (ICC=0.30). 51.6% of overall variability was due to predictors (Marginal R^2 =0.52) and 66% due to both predictors and intra-individual differences (Conditional R^2=0.66). These findings are mirrored in the unimputed data (appendix table 11.6 & 11.7).

Table 5.3:Unconditional and conditional growth model

		Unconditional Model			Fully conditional Model		
	Predictors	β (CI)	Std β (CI)	P value	B (CI)	Std (β)	P value
Fixed Effects	(Intercept)	42.32 41.58 to 43.06	0 (-0.05 to 0.05)	<0.001	23.31 (20.39 to 26.23)	0 (-0.03 to 0.03)	<0.001
	Measurement Occasion	0.19 0.05 to 0.33	0.04 (0.01 to 0.06)	0.008	-0.13 (-0.23 to -0.02)	-0.02 (-0.04 to -0.00)	0.023
	Age				0 (-0.03 to 0.03)	0 (-0.03 to 0.03)	0.951
	Sex				-1.31 (-2.12 to -0.50)	-0.05 (-0.09 to -0.02)	0.002
	HAQ				-2.59 (-3.10 to -2.08)	-0.18 (-0.21 to -0.14)	<0.001
	SF36MH				0.38 (0.35 to 0.41)	0.37 (0.35 to 0.40)	<0.001
	SF36BP				0.23 (0.19 to 0.26)	0.21 (0.18 to 0.25)	<0.001
	PGA				-0.04 (-0.05 to -0.03)	-0.08 (-0.11 to -0.06)	<0.001
Random Effects	σ^2	69.29			43.93		
	τ_{00}	70.55 $_{cenid}$			23.27 $_{cenid}$		
	τ_{11}	0.78 $_{cenid.\ visitx}$			0.27 $_{cenid.visitx}$		
	ρ_{01}	-0.35 $_{cenid}$			-0.60 $_{cenid}$		
	ICC	0.49			0.30		
	N	729 $_{cenid}$			729 $_{cenid}$		
	Marginal / Conditional R^2	0.001 / 0.488			0.516 / 0.660		

Intercept – Vitality values at baseline, β – Beta coefficient, std. - standardised, CI - Confidence Interval. σ2- Residual Variance, τ00 – Intercept variance, τ11 – Variance due to measurement occasion, ρ01– Covariance due to participant and measurement occasion, ICC – Intraclass correlation coefficient, N – Total number of participants, Vitality - SF36 Vitality, Age - Age at onset, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Pain - SF36 Bodily Pain, Mental health - SF36Mental health

5.5.3 Trajectories of Vitality

GBTM analysis of Vitality scores revealed that the two-group model best fit the data based on the fit statistics. Table 5.4 below shows the fit statistics for 1, 2 and 3 class models.

Table 5.4: Comparing fit statistics of 1, 2 and 3 group models.

No. of groups	**Fit Statistic**			
	BIC	AIC	LL	Entropy
1	-11,206.32	-11,197.13	-11,193.13	
2	-10,704.66	- 10,688.59	- 10,681.59	0.80
3	-10,580.26	- 10,555.01	-10,544.01	0.77

BIC – Bayesian Information Criteria, AIC – Akaike Information Criteria, LL – Log Likelihood.

The 2-group model had a lower BIC, AIC, -2LL and a higher entropy than the 3-group solution. Model parameters of the two-group model shown in table 5.5 show that the average posterior probability of each group (groupAPP) was 0.94 and 0.95, well over the recommended minimum value of 0.70. The odds of correct classification (Occ.pp) were over 5 for each group, indicating good assignment accuracy, further supporting the choice of the 2-group model. An alternative model specification using growth mixture modelling also favoured the 2-group membership (Appendix table 11.13).

Table 5.5: Fit statistics of each group in the 2-group model

2 Trajectory Groups – Classification Fit Statistics				
Groups	Count	groupAPP	occ_pp	TotProb
1	391	0.94	13.72	0.53
2	338	0.95	21.19	0.47

Count – Number of people in each group, groupAPP – Average posterior probability for each group, Occ-pp – Odds of correct classification based on weighted posterior probabilities, TotProb -the expected number based on the sums of the posterior probabilities.

These two groups were named the `Fatigue' and `No-Fatigue' groups, making up 53% and 47% of the total population, respectively. Table 5.6 below shows the parameter estimates of the groups, and figure 5.3 shows the progression of Vitality in the two groups over the study period and associated confidence intervals. Participants in the No-fatigue group had mean Vitality scores of 49.67 (CI:48.70 to 50.64) at baseline, and the net change over time was (ß = 1.70, CI:0.95 to 2.56). In the Fatigue group, the mean Vitality score at baseline was 35.23 (CI:34.45 to 36.00), and the rate of change over the study was (ß = 0.07, CI: -015 to 0.29). The mean baseline Vitality score in the fatigue group was greater than one standard deviation less than the average population levels.

Table 5.6: Trajectory estimates of two fatigue groups.

Groups	Parameters	Estimate	CI		Group membership (%)
Fatigue	Intercept	35.23	34.45	36.00*	52.65
	Linear	0.07	-0.15	0.29	
No Fatigue	Intercept	49.67	48.70	50.64*	47.35
	Linear	1.76	0.95	2.57*	
	Quadratic	-0.24	-0.37	-0.11*	

Intercept – Vitality values at baseline, Linear/Quadratic - Vitality growth rate per measurement occasion, * P ≤ 0.05
Group membership - % of total population assigned to each group, CI – 95% Confidence Interval

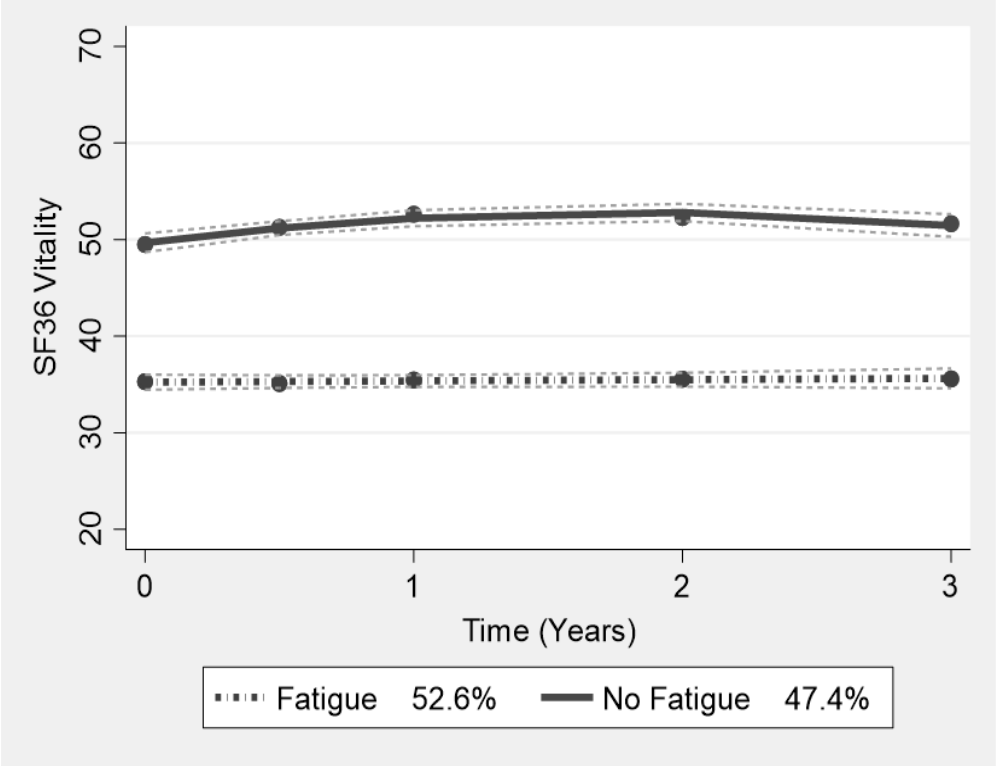


Figure 5.3: Progression of Vitality over the study period in 2 groups

Table 5.7 describes the baseline characteristics of the Vitality trajectory groups. Individuals in the Fatigue group were predominantly female (75% vs 60 %), had higher DAS28 (5.14 vs 4.39), worse pain and mental health, and worse HAQ and PGA than those in the No-fatigue group.

Table 5.8 showed that BMI, HAQ, mental health, bodily pain, female sex, and comorbidities were collectively associated with fatigue group membership in the multivariable analysis. In addition to the afore mentioned variables, DAS28ESR and component indices except SJC were associated with fatigue group membership in univariate analysis. These associations, however were not found in multivariable analysis

(please note that DAS28ESR was not included in multivariable modelling).

Table 5.7: Baseline Characteristics based on Trajectory groups.

Characteristic	Unit	Baseline	
		Group 1 (Fatigue n=391)	Group 2 (No-Fatigue n=338)
Demographics			
Age	years	56.44(13.06)	56.67 (14.00)
Sex	Female	292 (74.68)*	201 (59.47)*
White Ethnicity		384 (98.21)	325 (96.15)
BMI	Kg/m2	27.91 (24.75 to 31.63)~	26.29 (23.72 to 29.36)~
Smoking status			
Never		137(36.05)*	143 (43.73)*
Current		109 (28.68)*	78 (23.85)*
Ex-Smoker		134 (35.26)*	106 (32.42)*
Disease Activity/Inflammation			
DAS28ESR	0 - 9.4	5.14 (1.47)	4.39 (1.53)
TJC	0 - 28	8 (3 to 14)~	4 (1 to 9)~
SJC	0 - 28	5 (2 to 10)~	4 (1 to 9)~
Seropositive		211 (63.94)*	178 (61.38)*
Erosion		113 (30.87)*	100 (32.36)*
Nodules		36 (10.11)*	28 (9.21)*
Hb	Mg/dl	13.08 (1.40)	13.15 (1.46)
ESR	Mm/hr	27 (13 to 47)	20 (12 to 36)
PROMS			
Pain	0 - 100	29.21 (9.14)	37.90 (9.14)
Vitality	0 - 100	35.20 (8.51)	49.09 (8.39)
Mental health	0 - 100	42.33(10.84)	52.02 (9.15)
HAQ	0 - 3	1.29 (0.72)	0.79 (0.67)
PGA	0 - 100	50.74(23.60)	35.92 (24.09)
Comorbidity /Medication			
RDCI	0 – 9	1 (0 to 2)~	0 (0 to 1)~
medication		318 (81.33)	291 (86.09)
Others			
Disease Duration median, IQR	months	12 (7 to 23)~	13 (8 to 21)~

Values are mean(SD) unless otherwise stated.
Vitality - SF36 Vitality, Age - Age at onset, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Hb - haemoglobin Concentration, DAS28ESR - Disease Activity Score (ESR), Pain - SF36 Bodily Pain, Mental health - SF36Mental health, RDCI - Rheumatic Disease Comorbidity Index, Duration - Disease Duration in Months, TJC - Tender Joint Count, SJC - Swollen Joint Count, ESR - Erythrocyte Sedimentation rate, BMI - Body Mass Index, IQR - Inter Quartile Range. (*=n(%), ~=median(IQR))

Table 5.8: Baseline predictors of Vitality group membership

Variable	Univariate logistic regression		Multiple logistic regression
	Unit	OR (CI)	OR (CI)
Age	Yrs.	1.01 (0.99 to 1.02)	1.01 (0.99 to 1.02)
BMI	kg/m^2	1.07* (1.02 to1.11)	1.06* (1.02 to 1.10)
HAQ	0 -3	1.57* (1.09 to 2.27)	1.69* (1.22 to 2.33)
Mental health	0 - 100	0.94* (0.92 to 0.96	0.94* (0.92 to 0.96)
Bodily pain	0 -100	0.95* (0.93 to 0.98)	0.96* (0.93 to 0.98)
Female		1.71* (1.08 to 2.71)	1.94* (1.28 to 2.95)
Comorbidities (RDCI)	0 – 9	1.56* (1.20 to 1.54)	1.32* (1.12 to 1.56)
PGA	0 - 100	1.02*(1.01 to 1.03)	
TJC	0-28	1.08* (1.05 to 1.11)	
SJC	0 -28	1.02 (0.99 to 1.04)	
ESR	(mm/hr)	1.01* (1.00 to 1.02)	
DAS28-ESR		1.40* (1.25 to 1.56)	

* $p \leq 0.05$.

Univariate and multiple logistic regressions giving odds ratios (OR) with 95% confidence intervals (CI). Age – Age at onset, HAQ - Health Assessment Questionnaire, PGA - Patient's Global Assessment of Disease Activity, Hb - haemoglobin Concentration, DAS28ESR - Disease Activity Score (ESR), RDCI - Rheumatic Disease Comorbidity Index, TJC - Tender Joint Count, SJC - Swollen Joint Count, ESR - Erythrocyte Sedimentation rate, BMI - Body Mass Index

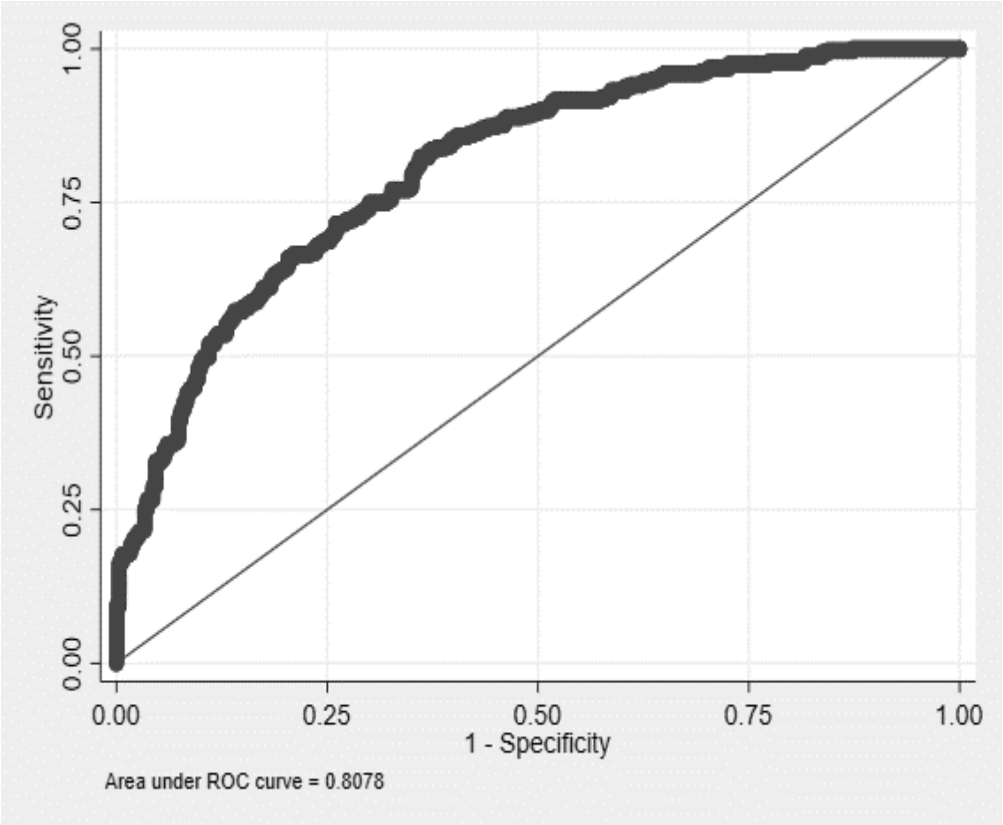


Figure 5.4: AUROC graph of logistic regression model describing baseline predictors of the fatigue- group membership.

5.5.4 Sensitivity Analysis

Growth analysis conducted on participants with a disease duration of six months or less and five complete Vitality measurements yielded similar results to those presented above. (Appendix tables 11.8 to 11.12). The linear growth model was selected, and the progression of fatigue over the study duration was comparable.

5.6 Discussion

5.6.1 Key findings and interpretations

This study examined trajectories of fatigue in a population of people with early rheumatoid arthritis. Key findings from this study were that fatigue remained persistent over time, undergoing no clinically meaningful change during the study period. Although baseline levels of pain, mental health, functional capacity, and female sex were associated with fatigue, these covariates were not associated with the shape of the fatigue trajectory. In the RA population, two groups of individuals exist based on fatigue trajectories – people with fatigue and people without fatigue. In addition, specific baseline characteristics, including female sex, pain, mental health, and functional ability, were associated with membership in the fatigue group. Unsurprisingly, markers of disease activity and inflammation such as SJC and ESR were not associated with fatigue in this study.

Our examination of the shape of the fatigue trajectory revealed that the linear trajectory best suited the data. This is not surprising given that fatigue levels did not exhibit clinically meaningful change over the study duration. Therefore, it is important to note that although a statistically significant rate of change was elicited in this study, this change may not be clinically relevant, as it is lower than the smallest score perceived by the patient – the MCID. This change value could be attributed to the measurement error of the fatigue instrument rather than an actual change in fatigue levels over the study period.

Our study aligns with other cohort studies on the population mean fatigue trajectory, finding fatigue to be persistent despite improvements in inflammation and disease activity. These studies, however, did not explore heterogeneity in fatigue trajectories(Feldthusen et al., 2016b). Although female sex and pain, HAQ, mental health and PGA at baseline were associated with the mean fatigue trajectory in our study, corroborating findings from several cross-sectional studies(Pollard et al., 2006, Belza et al., 1993, Van Hoogmoed et al., 2010); surprisingly, none of these factors were significantly associated with the progression of fatigue. The lack of association between fatigue and inflammatory markers, e.g., ESR and measures of disease activity, e.g., TJC and SJC, has also been reported by other studies(Hammer et al., 2021, Pollard et al., 2006). A positive association between PGA and fatigue could explain the relationship between fatigue and disease activity, measured with DAS28, reported by some studies, probably because PGA is a component of the DAS28 composite score. This study identified two distinct fatigue groups: people with no fatigue and those with persistent fatigue. This grouping appears to align with our dichotomy of Vitality in the previous chapter using standard deviations from the population normed mean. In contrast to our study, a study exploring the heterogeneity in the course of fatigue using a similar methodology reported a group with improving fatigue and additional groups in females(Druce et al., 2015b). I did not find this distinction in this study and in my study, trajectories of fatigue remained similar in both sexes. The said study reportedly focused on people with clinically relevant fatigue at

baseline. The lack of a universally acceptable fatigue definition and cut off levels preclude a head-to-head comparison with this study on the basis of the characteristics study population. However, the study included at least two fatigue measurements, while I had at least three, allowing for better ascertainment of sustained change over time in our study. In addition, my sensitivity analysis included five complete measurements and yielded similar results.

Another cohort study examining RA fatigue heterogeneity found three trajectory groups. (Provan et al., 2020). The study was based on a smaller sample consisting of people treated with biologic DMARDs. Although the study concluded that fatigue was associated with similar PROMs elicited by my study, modelling was conducted using a priori fatigue classifications in a population with established RA. I on the other hand, utilised a data-driven approach. A priori classifications may be constrained by our prior knowledge and do not allow the introduction of new ideas; a data-driven approach is more flexible and permits the introduction of novel ideas. This is one of the real benefits of novel machine learning approaches that have found recent use in medical research.

Female sex, in addition to patient-reported outcomes like mental health, pain, and functional capacity, were associated with fatigue in this study; findings which are consistent across other disease conditions such as inflammatory bowel disease, post-stroke and post-viral infections (Kjeverud et al., 2020, Klusmann et al., 2021). In addition, the persistence of fatigue despite adequate control of inflammatory disease or in the

presence of conditions with no demonstrated inflammatory component such as Fibromyalgia(Overman et al., 2016), and a failure to demonstrate an association between disease activity markers and fatigue, may indicate that the causal mechanisms of fatigue may be person and not disease-specific. Furthermore, it has been suggested in the knee osteoarthritis population that at least some sensitisation may be due to a trait (person-specific characteristic) or pathology (state)(Neogi et al., 2015).

Factors associated with fatigue in RA and non-RA disease conditions are components of a cluster of characteristics described in RA as fibromyalgianess. Fibromyalgianess is associated with central sensitisation, involving hyper-excitement of the central neurons characterised by amplification of noxious stimulus and increased sensitivity to environmental stimuli such as heat or light (Yunus, 2007). Central sensitisation is an established pain mechanism in musculoskeletal diseases, and evidence of an association with fatigue has been reported(Druce and McBeth, 2019). Central sensitisation could explain fatigue persistence and association with pain, mental health and lower functional capacity and overall worse outcomes even in the presence of well-controlled inflammatory disease(Wolfe, 2009, Lee et al., 2018).

It is noteworthy that randomised controlled trials (RCTs) of the effect of biologic therapies on RA showed a slight improvement in fatigue levels after six months of follow-up after treatment with biologic DMARD therapies. However, these studies did not examine the long-term

changes in fatigue (Almeida et al., 2016). Although it remains possible, fatigue in RA may originate from an inflammatory process, whereby pro-inflammatory cytokines fundamental to the development of an immune response trigger long-term changes in brain architecture, neural pathways, and sensitisation culminating in central sensitisation and fibromyalgianess (Korte and Straub, 2019); the persistence of inflammation does not seem to affect the progression of fatigue in this population.

RCTs have also demonstrated a therapeutic 'window of opportunity' in RA management associated with better outcomes in terms of radiological damage and disability (Chauffier et al., 2012). There is some debate about the duration of this therapeutic window; however, up to two years, post-diagnosis is supported by the literature. It is proposed that intervening while the disease process is less mature and more reversible would facilitate the modulation of the disease. This non-linear progression has been demonstrated in studies describing disease activity, pain, and psychological distress in RA (McWilliams et al., 2019, Norton et al., 2013). Although not designed to examine the effect of treatment, this study revealed that fatigue was present at the start of RA and underscored another consideration that the causes of fatigue may occur much earlier in the disease course, possibly even in the preclinical stages of the disease. More studies on populations at risk of developing RA may provide greater insight into the mechanisms of fatigue(Deane et al., 2010).

5.6.1.1 Strengths and limitations

This is a large longitudinal study with over 700 participants. Fatigue was modelled using robust statistical techniques that examined the trajectory of fatigue presented by the data. The inclusion of three or more measurements allowed the modelling of both linear and non-linear trajectories, and sensitivity analysis using individuals with complete measurements facilitated the validation of study findings.

This study is subject to several limitations. High attrition rates precluded the inclusion of additional timepoints in the longitudinal analysis. Another important consideration is the interval between measurements. Vitality was measured at least at a six-month interval but mostly annually. Perhaps a shorter measurement interval may provide more detail on the shape of the fatigue trajectory and factors associated with the trajectory. Seasonal variation in fatigue measurements did not show a significant association in this study.

RA treatment strategies have evolved since the inception of this cohort, and it is unclear if modelling these new treatment strategies would present a different population and provide additional trajectories. This study did not address some notable factors associated with fatigue (e.g., sleep quality). Univariate Group-based trajectory analysis was employed in data analysis; perhaps, modelling joint trajectories of other significant fatigue risk factors would provide more insight into the nature of heterogeneity observed in this analysis.

5.6.2 Conclusion

Fatigue is a persistent problem in people with RA, remaining stable through the disease course. Diverse baseline characteristics that may indicate central mechanisms are associated with persistent fatigue. Groups that may develop persistent fatigue may be identified early in the disease process. These individuals may be candidates for interventions to reduce fatigue, in addition to disease-modifying interventions.

5.6.3 Summary

This chapter concludes the investigation of fatigue in early RA. I have shown that fatigue is prevalent and persistent in RA. Patient-reported outcomes rather than disease activity and inflammation were associated with fatigue. Individuals with baseline characteristics that may be indicative of central sensitisation were found to be associated with fatigue. These traits could explain the similarities in the progression of fatigue in other disease conditions and the relationship between fatigue and other characteristics associated with central sensitisation, e.g., pain. I posit that individuals who possess some of these traits may experience more fatigue than people without these traits in the absence of a specific disease condition.

In the next chapter, I investigate fatigue in a general population sample and examine the association between fatigue and central sensitisation in people with knee pain and without knee pain.

6. Fatigue in the general population – a knee pain in the community study

6.1 Outline

In the previous chapters, I demonstrated that fatigue was prevalent and persistent in people with early RA and that people who report fatigue at baseline continue to experience fatigue. I think that this may be related to traits associated with central sensitisation. Consequently, I hypothesise that people with specific characteristics were more likely to report fatigue. In this chapter, I explore these characteristics associated with fatigue in the general population. We investigate differences in people with and without knee pain and assess the relationship between fatigue and central sensitisation.

This chapter extends findings from the previous chapter and explores evidence around the determinants of fatigue in the general population. The research described in this chapter will provide greater insight into factors associated with fatigue, considering features of central sensitisation.

6.2 Introduction

Although fatigue is a physiologic experience, it can have a profound impact on the quality of life of some individuals. Studies investigating the prevalence of fatigue in the general population reveal prevalence ranging between 2 – 59%, depending on the study population, country of study, duration of fatigue assessed and the definition of fatigue (Engberg et al., 2017, Watt et al., 2000, Schwarz et al., 2003, Cullen et al., 2002, Ridsdale

et al., 1993, Nijrolder et al., 2009, Ricci et al., 2007). Fatigue is one of the commonest symptoms reported in primary care, accounting for about 20% of all general practice consultations(Cullen et al., 2002). In a general population sample, fatigue was found to be associated with excess mortality (Basu et al., 2016).

Characteristics associated with fatigue bisect several disease conditions, including musculoskeletal diseases, cancer, cardiovascular, endocrine, gastrointestinal, haematologic and respiratory conditions. Longitudinal studies in multiple disease conditions paint a picture of persistent fatigue despite objective improvements in the underlying disease process(Bower, 2014, Nijrolder et al., 2009). Traits associated with abnormal central processing of sensory stimuli have been found to be associated with fatigue and predict worse outcomes in many disease conditions (Neogi et al., 2015, van Steenbergen et al., 2015, Téllez et al., 2006).

Central sensitisation can be characterised by the amplification of noxious and non-noxious peripheral signals. It results in amplified responses to pain and heightened sensitivity to physical and environmental stimuli such as heat, sound and light. In addition, it could explain the discordant relationship between persistent fatigue and an underlying disease process(Yunus, 2007).

Several chronic disease conditions are characterised by central sensitisation, including chronic fatigue syndrome, irritable bowel syndrome, fibromyalgia, and osteoarthritis (Midenfjord et al., 2021,

Williams, 2018, Bourke et al., 2021). Traits associated with central sensitisation have been identified in people with knee pain (Akin-Akinyosoye et al., 2018). Identifying people with these traits could represent targets for interventions to correct central processing abnormalities and improve the quality of life of people with fatigue.

Most general population studies focus on demographic and social determinants, e.g., age, sex, and socioeconomic status. Other factors associated with fatigue in many disease conditions, such as pain, mental health, functional capacity, and sleep disturbance, are less represented in general population studies. Consequently, the relationship between fatigue and central sensitisation has not been fully elucidated in the general population.

6.3 Aims and Objectives

6.3.1 Aim

This study aims to investigate fatigue in the general population.

6.3.2 Objectives

- Describe the prevalence of fatigue in the general population.
- Examine characteristics associated with fatigue and persistent fatigue.
- Investigate the relationship between fatigue and central sensitisation.

6.3.3 Hypothesis

I hypothesise that traits associated with central sensitisation are associated with fatigue in the general population.

6.4 Methods

6.4.1 Study design/source of data

Data were from the Knee Pain and related health In the Community (KPIC), a UK community-based cohort of adults in the East Midlands, United Kingdom. KPIC study design was described in the chapter three.

6.4.2 Exposures and outcomes

6.4.2.1 Outcome variable

The outcome variable of this study was the Vitality subscale of the SF12 questionnaire (SF12Vitality), normed to the UK general population on a scale of 0 – 100, where higher values represented a better quality of life. (Ware et al., 1996).

A secondary study outcome was a binary (y/n) classification of fatigue, based on the question "***In the past month, did you feel tired on most days?"*** on a 5-point Likert scale ranging from never to always (never, sometimes, often and always). Participants who answered '***often'*** and '***always'*** were categorised as the fatigue group.

6.4.2.2 Exposure variables

Demographic characteristics including self-reported age (in years), sex (male/female), body mass index (kg/m^2), were assessed. Mental Health

was measured using the Hospital Anxiety and Depression scale (HADs). Sleep was assessed using the sleep adequacy dimension of the Medical Outcomes Survey (MOS) sleep. Pain was assessed with the bodily pain subscale SF12, normed to the UK population. Comorbidities were measured as a count of self-reported comorbidities, ranging from 0 (no comorbidities) to 10 (the highest number of comorbidities reported by a single participant in the dataset).

In the KPIC sub-population that had further assessments, inflammation was measured by C-Reactive Protein (CRP). Pressure pain detection threshold (PPT) measured in the anterior tibia and sternum were used as proxy measures of central sensitisation.

6.4.3 Statistical Analysis

The study population was described with means and standard deviations and median and inter-quartile ranges for normally and non-normally distributed continuous variables, respectively. Normality was assessed graphically with histograms. Categorical variables were described using frequencies and proportions.

SF12Vitality was dichotomised based on the population mean (using a similar process described for SF36 in chapter 4). The crude prevalence of fatigue in KPIC was calculated by dividing the number of people classified as having fatigue by the total number of people in the total KPIC population.

Fatigue prevalence was standardized by age in 5-year bands and sex to the European standard population (ESP) 2013. 95% confidence intervals

were derived based on the assumption of a gamma distribution. Prevalence was estimated using the 'Epitools' package for the R statistical software (Aragon, 2020).

The association between baseline covariates and fatigue was examined using scatter plots, and Spearman correlation coefficients (r), with Bonferroni correction to account for multiple testing. In order to investigate factors associated with fatigue in year one, adjustment was made for baseline SF12Vitality in the regression model(Vickers and Altman, 2001). Bivariate and multivariable linear regression with robust standard errors was conducted, and standardized regression coefficients estimated. Linear regression diagnostics were examined graphically. Multicollinearity was investigated using the variance inflation factor (Chapter 3). Subgroup analysis was conducted on participants with and without knee pain.

The association between SF12Vitality and central sensitisation was assessed using correlation, and linear regression analysis, with SF12Vitality as the outcome and PPT as the main exposure variable.

Secondary analysis was performed using a y/n fatigue classification as the outcome variable. In this case, logistic regression, accounting for clustering by GP practice, was conducted due to issues with convergence of log-binomial models used to estimate risk ratios (chapter 3).

In order to generalise study findings to the general population, given KPIC sampling framework, the following analysis was conducted. Firstly, the total KPIC population distribution on age and sex were selected and

compared to the UK East Midlands census populations (Appendix 11.14). The results showed that the 60-69 age group, and females were slightly overrepresented in the KPIC population. Weights were then applied to the KPIC population using raking (chapter 3) and presented in appendix tables 11.15 to 11.18.

Statistical analysis was conducted in STATA® version 17(Stata, 2021) and R (R Core Team, 2021) – Epitools(Aragon, 2020) and Anesrake (Pasek and Pasek, 2018)packages. All estimates were presented with 95% confidence intervals (CIs), and unless otherwise stated, significance levels were set to 0.05.

6.5 Results

6.5.1 Overview of the study population

The KPIC study dataset included entries for 9513 participants at baseline. Of this number, 9,406 (98.88%) reported SF12Vitality at baseline. In year one, 4,734 participants (49.76% of the total KPIC population) remained in the study. Of this number, 4,630 (97.80%) reported SF12Vitality. A total of 420 (4.42% of the total KPIC population) participants provided further clinical assessment at baseline (PPT, CRP) data. Of these, 322 reported knee pain at baseline; of this number, 219 were classified in the early knee pain group. Figure 6.1 shows a flow diagram of KPIC participants used in this study.

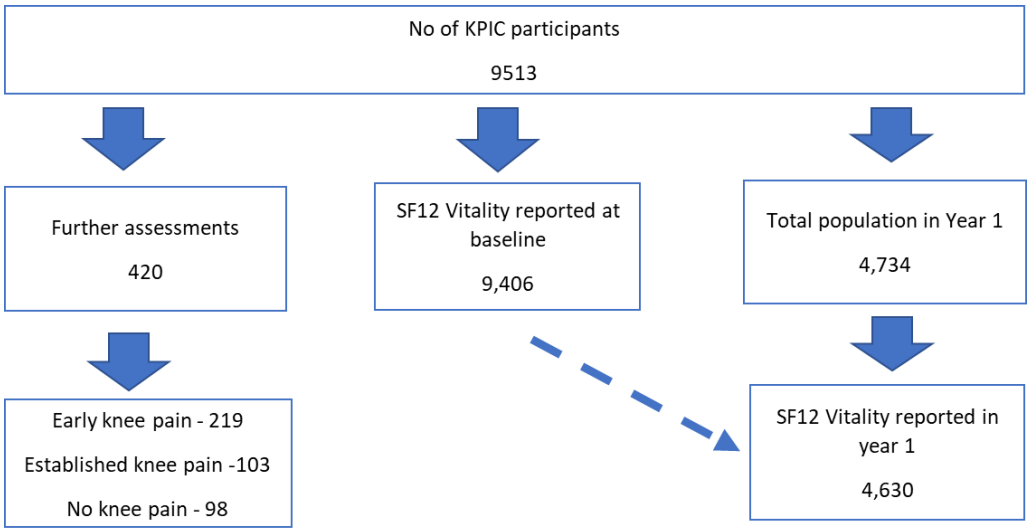


Figure 6.1: Flowchart describing KPIC dataset

Table 6.1 shows that the mean age of the population was 62.1(SD:10.6) years, 57% were female, and the median BMI was 26.5 (IQR: 23.76 to 29.70). 63.4% of the population reported fatigue, and the mean SF12Vitality score was 43.39 (SD:11.5). The mean bodily pain score was 41(SD:10). 45% of the sample reported knee pain. The median HADs Anxiety score was 6 (IQR:3 to 9), while that of HADs depression was 3 (IQR:2 to 6). Cardiovascular comorbidities were the commonest in this population (43.4%). 1.4% of the population reported chronic fatigue syndrome, and 2.6% reported a diagnosis of fibromyalgia.

Table 6.1: Overview of total KPIC population at baseline

Characteristics	Unit	Values	n
Age	years	62.09 (10.58)	9,452
BMI	kg/m2	26.45 (23.76 to 29.70) *	9,070
Female		56.74%	5,365
Fatigue		63.36 %	5,944
SF12Vitality	0 -100	43.39 (11.51)	9,406
Pain	0-100	41.00 (10.37)	9,380
Knee pain		45.3%	2,686
HADs- Anxiety	0-21	6 (3 to 9) *	9,418
HADs- Depression	0 - 21	3 (2 to 6) *	9,422
Sleep Adequacy[1]	0-100	57.50 (29.54)	9,370
Comorbidity	count (0-10)	1 (0 to 2) *	9,513

Comorbidity(details)	%	n
CVS	43.4	4,127
COPD	1.05	100
Asthma	3.54	337
IPF	0.02	2
Thyroid Disease	3.85	366
Diabetes	9.46	900
Liver & Gall Bladder Didease	0.81	77
Kidney Dx	0.7	67
MS	0.09	9
Stroke	2.85	271
Rheumatoid Arthritis	0.9	86
SLE	0.14	13
Psoriatic Arthritis	0.34	32
Ankylosing Spondylitis	0.19	18
Cancer	8.71	828
Fibromyalgia	2.59	246
Gout	0.99	94
knee/hip OA	17.3	1,645
IBS	10.13	963
CFS	1.4	133

Values are mean (SD) unless otherwise stated. * - median (IQR)
CRP - C-Reactive Protein, CVS – Cardiovascular Disease, COPD – Chronic Obstructive Pulmonary Disease, IPF – Idiopathic Pulmonary Fibrosis, Dx – Disease, MS – Multiple Sclerosis, SLE – Systemic Lupus Erythematosus, OA – Osteoarthritis, IBS – Irritable Bowel Syndrome, CFS – Chronic Fatigue Syndrome, Y - Yes

[1] Higher scores on the sleep adequacy scale represent more adequate sleep.

Table 6.2 shows an overview of the descriptive characteristics of the population with knee pain compared with those who did not report knee pain. These sub-populations did not statistically significantly differ in terms of age and most comorbidities assessed. Conversely, the proportion of females was slightly higher in the no-knee pain group (p=0.044). People with knee pain recorded significantly lower SF12Vitality, bodily pain, and sleep adequacy scores (38.62 (SD:11.45), 33.66(SD:10.26) and 56.56 (SD:30.00), respectively – each indicating worse quality of life), and higher anxiety and depression scores (7(IQR: 4 to 11) and 5(IQR:3 to 9) respectively). People with knee pain also reported a significantly higher proportion of cardiovascular comorbidities, Fibromyalgia, knee/hip OA, IBS and CFS ($p<0.0001$). In addition, RA and diabetes occurred more commonly in the knee pain group($p<0.001$). Cardiovascular comorbidities were the commonest reported comorbidities in both populations.

Table 6.2: Descriptive Characteristics of the knee and no-knee pain KPIC subpopulations at baseline

Characteristics	Unit	Knee pain (n - 2686)	No-Knee pain (n - 6,827)	P value
Demographics				
Age	years	62.09 (10.62)	62.11 (10.65)	0.6597
BMI	kg/m2	25.60 (23.24 to 28.52) ~ ~	25.92 (23.41 to 28.97) ~	<0.0001
Female		1,557(58.38) *	3,808 (55.78) *	0.044
Fatigue				
Fatigue		2,109 (79.80) *	3,835 (56.92) *	<0.0001
SF12Vitality	0 -100	38.62(11.45)	45.25 (10.99)	<0.0001
Pain				
SF12 Bodily Pain	0-100	33.66 (10.26)	43.89 (8.87)	<0.0001
Mental Health				
Anxiety	0-21	7 (4 to 11) ~	5 (3 to 8) ~	<0.0001
Depression	0 - 21	5 (3 to 9) ~	3 (1 to 5) ~	<0.0001
Sleep				
Sleep Adequacy	0-100	56.56 (30.00)	61.19 (28.38)	<0.0001
Comorbidity	count (0-10)	1 (0 to 2) ~	1 (0 to 1) ~	<0.0001
*Comorbidity(details)		**%**	**%**	**P value**
CVS		1,332 (49.59)	2,795 (40.96)	<0.0001
COPD		34 (1.27)	66 (0.97)	0.198
Asthma		105 (3.95)	231 (3.38)	0.181
IPF		-	2 (0.03)	-
Thyroid Disease		107 (3.98)	259 (3.79)	0.665
Diabetes		350 (13.03)	550 (8.06)	<0.001
Liver & Gall Bladder Dx		22 (0.82)	55 (0.81)	0.948
Kidney Dx		23 (0.86)	44 (0.64)	0.266
MS		3 (0.11)	6 (0.09)	0.734
Stroke		97(3.61)	174 (2.55)	0.005
Rheumatoid Arthritis		38(1.41)	48 (0.70)	<0.001
SLE		6 (0.22)	7 (0.10)	0.151
Psoriatic Arthritis		12 (0.45)	20 (0.29)	0.244
Ankylosing Spondylitis		6 (0.22)	12 (0.18)	0.631
Cancer		215 (8.00)	613 (8.98)	0.128
Fibromyalgia		145 (5.4)	101 (1.48)	<0.0001
Gout		28 (1.04)	99 (0.97)	0.737
knee/hip OA		1,003 (37.34)	642 (9.41)	<0.0001
IBS		963 (10.13)	597 (8.75)	<0.0001
CFS		60 (2.23)	73 (1.07)	<0.0001

Values are mean **(**SD) unless stated otherwise

PPT – Pressure Pain Threshold, CRP- C-Reactive Protein, CVS – Cardiovascular Disease, COPD – Chronic Obstructive Pulmonary Disease, Dx – Disease, MS – Multiple Sclerosis, SLE – Systemic Lupus Erythematosus, OA – Osteoarthritis, IBS – Irritable Bowel Syndrome, CFS – Chronic Fatigue Syndrome, Y – Yes,(*=n (%) , ~=median(IQR))

6.5.2 Prevalence of fatigue in the general population

About half to two-thirds of the total KPIC population reported SF12Vitality values greater than one SD below the normative UK population mean. The crude prevalence obtained was 43%, and the age-sex standardised prevalence was 45% (CI:35% to 64%).

In the knee pain sub-population, crude fatigue prevalence was 62%, and age-sex standardised prevalence 70% (CI: 50% to 100%). In the no-knee pain population, however, the crude prevalence was 36%, and the age-sex standardised prevalence was 31% (CI: 29% to 56%).

6.5.3 Risk factors for fatigue at baseline.

In the pairwise correlation matrix of baseline covariates shown in table 6.3, SF12Vitality exhibited a statistically significant strong negative correlation with HADs depression (r = -0.7) and a moderate negative correlation with HADs anxiety (r =-0.5), i.e., lower vitality was associated with greater depression and anxiety. A moderate positive correlation was shown between SF12Vitality and bodily pain score and sleep (r = 0.5 and r = 0.6, respectively), meaning that more vitality was associated with less pain and better sleep. The significant positive correlations between anxiety and depression (r = 0.6) and the negative correlation between bodily pain score and HADs-depression score (r = -0.5) are noteworthy from the table. In summary, higher pain, poor sleep, and higher anxiety and depression symptoms were correlated with fatigue (lower vitality). This pattern was mirrored in both the knee pain and no-knee pain sub-populations (tables 6.4 and 6.5).

Table 6.3: Correlation matrix between SF12Vitality and covariates at baseline in the total KPIC population

Characteristics	Age	BMI	Bodily Pain	Depression	Anxiety	Sleep	Comorbidity	Sex	SF12 Vitality
Age	1								
BMI	-0.0165	1							
Pain	-0.0513*	-0.2200*	1						
Depression	-0.0038	0.2133*	-0.5239*	1					
Anxiety	-0.1636*	0.0716*	-0.3775*	0.5915*	1				
Sleep	0.1609*	-0.1097*	0.3589*	-0.4330*	-0.4340*	1			
Comorbidity	0.3511*	0.2086*	-0.3132*	0.2251*	0.1073*	-0.1072*	1		
Sex	-0.0658*	-0.0685*	-0.0649*	-0.0114	0.1741*	-0.1262*	-0.0086	1	
SF12 Vitality	0.0094	-0.2139*	0.5228*	-0.6545*	-0.4671*	0.4529*	-0.2593*	-0.0789*	1

*P≤0.05, Values represent Spearman's correlation coefficient, rho(r)

Table 6.4: Correlation matrix between SF12Vitality and covariates at baseline in people with knee pain (n=2686)

knee pain (n=2444)									
	Age	BMI	Bodily Pain	Depression	Anxiety	Sleep	Comorbidity	Sex	SF12Vitality
Age	1								
BMI	-0.0481	1							
Bodily pain	-0.0548	-0.2442*	1						
Depression	-0.0678*	0.2170*	-0.5686*	1					
Anxiety	-0.1704*	0.0987*	-0.4327*	0.6459*	1				
Sleep score	0.1715*	-0.1126*	0.3570*	-0.4451*	-0.4236*	1			
Comorbidity	0.3523*	0.2397*	-0.3040*	0.2140*	0.1203*	-0.0937*	1		
Sex	-0.0456	-0.0359	-0.0243	-0.014	0.1182*	-0.1196*	0.0243	1	
SF12Vitality	0.0352	-0.2333*	0.5329*	-0.6359*	-0.4746*	0.4526*	-0.2555*	-0.0795*	1
*P≤0.05, Values represent Spearman's correlation coefficient, rho(r)									

Table 6.5: Correlation matrix between SF12Vitality and covariates at baseline in people with no-knee pain (n=6827)

No-Knee pain (n=6320)									
	Age	BMI	Pain	Depression	Anxiety	Sleep	Comorbidity	Sex	SF12Vitality
Age	1								
BMI	-0.0058	1							
Bodily pain	-0.0593*	-0.1229*	1						
Depression	0.0192	0.1682*	-0.4269*	1					
Anxiety	-0.1641*	0.0044	-0.3022*	0.5402*	1				
Sleep score	0.1636*	-0.0728*	0.2952*	-0.3853*	-0.4070*	1			
Comorbidity	0.3610*	0.1834*	-0.2308*	0.1645*	0.0508*	-0.0584*	1		
Sex	-0.0730*	-0.1011*	-0.0732*	-0.0219	0.1945*	-0.1268*	-0.0271	1	
Vitality	0.002	-0.1614*	0.4496*	-0.6205*	-0.4271*	0.4139*	-0.2022*	-0.0754*	1
*P≤0.05, Values represent spearman's correlation coefficient, rho(r)									

Table 6.6:Bivariate linear regression analysis describing the association between SF12Vitality and covariates in the total KPIC population at baseline.

Characteristics	unit/range	ß	CI	p value	n
Age	Years	0.01	-0.02 to 0.03	0.661	9,350
BMI	kg/m2	-0.52	-0.57 to -0.48	<0.001	8,976
Sex	Female	-1.95	-2.41 to -1.48	<0.001	9,355
Bodily pain	0-100	0.60	0.58 to 0.62	<0.001	9,342
Anxiety	0 -21	-1.32	-1.37 to -1.27	<0.001	9,350
Depression	0-21	-2.03	-2.08 to -1.99	<0.001	9,354
Sleep score	0-100	0.18	0 .17 to 0.20	<0.001	9,303
Comorbidity	count (0-10)	-2.99	-3.20 to -2.77	<0.001	9,406
ß = Regression coefficient, CI =95% Confidence interval					

In bivariate analysis BMI, sex, bodily pain score, anxiety, depression, sleep score, and comorbidities were significantly associated with SF12Vitality. Female sex, BMI, depression, and anxiety were inversely associated with SF12Vitality at baseline, while better sleep and pain scores were directly associated with SF12Vitality. Age was not significantly associated with SF12Vitality. Bivariate linear regression analysis of the knee pain and no-knee pain subpopulations showed a similar pattern of associations with fatigue, table 6.7.

Table 6.7: Bivariate regression models describing the association between SF12Vitality and covariates in the knee and no-knee pain subpopulations at baseline.

Characteristics	Unit	Knee pain				No-knee pain			
		ß	CI	p value	n	ß	CI	P value	n
Age	years	0.02	-0.024 to 0.06	0.404	2,633	-0.002	-0.03 to 0.02	0.892	6,717
BMI	kg/m2	-0.44	-0.51 to -0.37	<0.001	2,506	-0.41	-0.47 to -0.36	<0.001	6,470
Female		-1.83	-2.71 to -0.95	<0.001	2,633	-1.81	-2.33 to -1.29	<0.001	6,722
Bodily pain	0 - 100	0.58	0.55 to 0.62	<0.001	2,635	0.59	0.56 to 0.62	<0.001	6,707
Anxiety	0 - 21	-1.18	-1.26 to -1.09	<0.001	2,639	-1.24	-1.30 to -1.17	<0.001	6,711
Depression	0 - 21	-1.67	-1.76 to -1.58	<0.001	2,639	-2.16	-2.23 to -2.08	<0.001	6,715
Sleep score	0-100	0.17	0.15 to 0.18	<0.001	2,616	0.16	0.16 to 0.17	<0.001	6,687
Comorbidity	count (0 -10)	-2.5	-2.85 to -2.14	<0.001	2,651	-2.49	-2.78 to -2.20	<0.001	6,755

Y – Yes, ß = regression coefficient, std = Standardised, CI – 95% confidence interval

Female sex, bodily pain score, sleep, anxiety, depression, BMI, and comorbidities were significantly associated with SF12Vitality, while age was not significantly associated with SF12Vitality in the multivariable analysis, table 6.8. Collectively, these factors explained about half of the variability of SF12Vitality in the study population (R^2 = 0.50). Depression contributed the most significant proportion to the model, with the highest standardised coefficient (std ß=-0.41). Linear regression assumptions were not violated, appendix figure 11.11, and multicollinearity was not detected (Mean VIF = 1.48).

Table 6.8:Multivariable linear regression showing the characteristics associated with fatigue in the total KPIC population.

	n = 8,777		R^2 = 0.50		
Characteristics	Unit	ß	CI	p value	Std ß
Bodily pain	0 - 100	0.22	0.20 to 0.24	<0.001	0.20
Depression	0 - 21	-1.30	-1.41 to -1.25	<0.001	-0.42
Anxiety	0 - 21	-0.11	-0.15 to -0.03	0.003	-0.03
Sleep score	0 - 100	0.07	0.06 to 0.07	<0.001	0.17
BMI	Kg/M^2	-0.12	-0.16 to -0.06	<0.001	-0.06
Comorbidities	0 - 7	-0.92	-1.10 to -0.64	<0.001	-0.08
Female		-1.05	-1.40 to -0.70	<0.001	-0.05

All $p \leq 0.05$, ß – Regression coefficient, Std ß = Standardized regression coefficient, CI – 95% Confidence Interval

Multivariable linear regression analysis describing the associations between fatigue and covariates in the knee and no-knee pain subpopulations at baseline are shown in table 6.9. Pain, anxiety, depression, sleep, comorbidities, and sex collectively explained 46% (R^2 = 0.46) and 48% (R^2 = 0.48) of fatigue variability in the knee and no-knee pain subpopulations, respectively. Analogous to the total population, depression was the dominant covariate associated with fatigue, std ß = -0.36 and -0.43 in the knee and no-knee pain subgroups, respectively.

Table 6.9: Multivariable linear regression model showing the characteristics associated with fatigue in the population with and without knee pain.

		Knee Pain, n= 2,447, R^2=0.46		No knee pain, n= 6,330, R^2=0.48	
Characteristics	Unit	ß (CI)	Std ß	ß (CI)	Std ß
Pain	0 - 100	0.21 (0.17 to 0.26)	0.19	0.23 (0.20 to 0.26)	0.18
Anxiety	0 - 21	-0.12 (-0.23 to -0.01)	-0.05	-0.12 (-0.20 to -0.05)	-0.04
Depression	0 - 21	-0.98 (-1.13 to -0.84)	-0.36	-1.50 (-1.59 to -1.40)	-0.43
Sleep	0 - 100	0.07 (0.06 to 0.09)	0.19	0.06 (0.05 to 0.07)	0.16
Female		-1.05 (-1.75 to -0.37)	-0.07	-1.06(-1.46 to -0.66)	-0.05
BMI	Kg/M^2	-0.12 (-0.18 to -0.06)	-0.07	-0.14 (-0.18 to -0.10)	-0.06
Comorbidities	0 - 10	-0.75 (-1.09 to 0.41)	-0.08	-0.87 (-1.10 to -0.63)	-0.08
All $p \leq 0.05$, ß = regression coefficient, std = Standardised, CI – 95% Confidence Interval					

6.5.4 Baseline factors associated with fatigue in year 1

The KPIC population in year 1 was similar to the starting population at baseline in terms of age, sex, SF12Vitality, bodily pain, fatigue, anxiety and depression. However, the proportion of people with knee pain was higher in the baseline population (45.30%) compared with the KPIC population in year 1 (37.20%), and the baseline population had better sleep adequacy scores (57.50, SD:29.54) compared with year 1 (52.48, SD:29.99) table 6.10.

Table 6.10: Characteristics of the KPIC population in year 1

Characteristics	**unit**	**Year 1**	**Baseline**
Demographics			
Age	years	63.95 (10.13)	62.09 (10.58)
Sex	female	2,728 (57.85) *	5,365 (56.74) *
Fatigue			
Fatigue	yes	2,942 (63.49) *	5,944 (63.36) *
SF12Vitality	0 -100	42.18 (11.29)	43.39(11.51)
Pain			
SF12 Bodily pain	0-100	40.67 (10.26)	41.00 (10.37)
Knee Pain	yes	1,729 (37.18)	2,686 (45.3)
Mental Health			
HADs- Anxiety	0 -21	5 (3 to 8) ~	6 (3 to 9) ~
HADs- Depression	0-21	3 (3 to 6) ~	3 (2 to 6) ~
Sleep			
Sleep Adequacy	0-100	52.48 (29.99)	57.50 (29.54)

*n (%), ~median (IQR), Baseline and year 1 values are means **(**SD) unless otherwise stated

Table 6.11 compares the baseline characteristics between the study population in year one and the population with baseline data only. The proportion of females was slightly lower in the baseline-only group (59.64 vs 57.85 (p=0.031). The baseline-only group had slightly higher SF12Vitality scores (42.98, SD:11.89 vs 43.80, SD:11.10, p=0.0005) and a more significant proportion of people reporting fatigue 64.38% vs 62.35%, p =0.041) at baseline. The proportion of people reporting knee pain at baseline was significantly higher in the year one group compared with the baseline-only group (48.26% vs 42.16%, p ≤0.001). The baseline-only group reported significantly less anxiety (3 IQR:3 vs 6, IQR:6, p =0.0275) but higher depression scores (4 IQR:5 vs 3 IQR:5, p≤0.0001) than the year one group.

Table 6.11: Comparing baseline characteristics between missing and non-missing KPIC participants in year 1.

Characteristics	Unit/range	Baseline & year 1	Baseline only	P value
Demographics				
Age	years	62.46 (10.1)	61.71(11)	0.0006
BMI	kg/m^2	26.36 (23.80 to 29.81) ~	26.57 (23.77 to 29.60) ~	0.2108
Sex	female	2,728 (57.85) *	2,637 (55.64) *	0.031
Fatigue				
Fatigue	yes	2,926 (62.35) *	3,018 (64.38) *	0.041
SF12Vitality	0 -100	43.80 (11.10)	42.98 (11.89)	0.0005
Pain				
SF12 Bodily Pain	0-100	41.04 (10.37)	40.95 (10.08)	0.6717
Knee Pain	yes	1,471 (48.26) *	1,215 (42.16) *	<0.001
Mental Health				
HADs- Anxiety median (IQR)	0 -21	6 (3 to 9) ~	3 (5 to 8) ~	0.0275
HADs- Depression	0-21	3 (1 to 6) ~	4 (2 to 7) ~	<0.0001
Sleep				
Sleep Adequacy	0-100	57.44 (28.98)	57.57 (30.10)	0.8270
Comorbidity	count (0-7)	1 (0 to 2) ~	1 (0 to 2) ~	0.0551

*- n (%), ~ - median (IQR), Baseline and year 1 values are means(SD) unless otherwise stated

Table 6.12 shows the paired correlation between SF12Vitality in year one and baseline covariates in the KPIC population. Baseline anxiety, depression, bodily pain, sleep scores, and BMI were significantly associated with SF12Vitality in year 1. Lower pain and better sleep showed a significant correlation with SF12Vitality ($r = 0.46$ and $r = 0.43$, $p \leq 0.05$ respectively), while there was a significant correlation between baseline depression ($r = -0.56$, $p \leq 0.05$), BMI ($r = -0.24$, $p \leq 0.05$) and female sex ($r = -0.11$) with SF12Vitality in year one. Unsurprisingly, SF12Vitality at baseline was significantly positively associated with SF12Vitality in year 1 ($r=0.7$, $p \leq 0.05$). Please note that higher values of bodily pain, SF12Vitality and sleep scores on the table denote a better quality of life. Correlation patterns were similar in the knee and no-knee pain sub-populations, tables 6.12 and 6.13.

Table 6.12: Correlation of baseline covariates and year 1 SF12Vitality in the total KPIC population

	SF12Vitality(yr1)
SF12Vitality (B)	0.6780*
Sleep	0.4294*
Age	0.0379
BMI	-0.2385*
Pain	0.4567*
Anxiety	-0.4097*
Depression	-0.5617*
Comorbidities	-0.2264*
Female	-0.1098*
*P ≤ 0.05, Values are spearman's rho (r),(B) - Baseline, (Yr1) - Year 1,	

Table 6.13: Correlation describing the relationships among baseline characteristics and SF12Vitality in year 1 in people with knee pain and no-knee pain.

	SF12Vitality (Yr1)	
	Knee pain	No-knee pain
SF12Vitality (B)	0.6585*	0.6557*
Sleep	0.4026*	0.4299*
Age	0.0235	0.0956*
BMI	-0.1974*	-0.2107*
Pain	0.3904*	0.4397*
Anxiety	-0.3705*	-0.3899*
Depression	-0.5298*	-0.5464*
Values are spearman's rho (r), * p ≤ 0.05, (B) - Baseline, (Yr1) -Year 1		

As expected from the correlation analysis, in the bivariate linear regression analysis, with year one SF12Vitality as the outcome variable and baseline characteristics as exposure variables - bodily pain, anxiety, depression, sleep scores, BMI, comorbidities, and female sex were significantly associated with SF12Vitality in year one in the total KPIC population. These baseline characteristics were also significantly associated with year one SF12Vitality scores in the knee pain and no-knee pain subpopulations, table 6.14 and 6.15.

Table 6.14:Bivariate associations between SF12Vitality in year one and baseline covariates in the total KPIC population.

Characteristics	Unit	ß	CI	p-value	n
Pain	0 - 100	0.53	0.50 to 0.55	<0.001	4,595
Anxiety	0 - 21	-1.16	-1.24 to -1.09	<0.001	4,604
Depression	0 - 21	-1.82	-1.90 to -1.75	<0.001	4,604
Sleep	0 - 100	0.17	0.16 to 0.18	<0.001	4,595
BMI	Kg/M^2	-0.08	-0.60 to -0.47	<0.001	4,439
Comorbidities	0 - 10	-2.46	-2.77 to -2.16	<0.001	4,630
Female		-2.51	-3.16 to -1.86	<0.001	4,612
ß – Regression coefficient, CI – 95% Confidence Interval					

Table 6.15: Bivariate associations between baseline characteristics and SF12Vitality in year 1 in the knee and no-knee pain sub-populations

Characteristics	unit	knee pain				No-knee pain			
		ß	CI	P value	n	ß	CI	P value	n
Age	years	0.08	0.03 to 0.14	0.004	1,435	0.04	0.001 to 0.08	0.047	3,192
Sex	Female	-2.30	-3.47 to -1.13	<0.001	1,429	-2.36	-3.12 to -1.61	<0.001	3,183
BMI	kg/m2	-0.38	-0.48 to -0.29	<0.001	1,361	-0.49	-0.57 to -0.41	<0.001	3,078
Pain	0-100	0.49	0.43 to 0.54	<0.001	1,429	0.52	0.48 to 0.56	<0.001	3,166
Anxiety	0 -21	-0.99	-1.11 to -0.87	<0.001	1,430	-1.11	-1.21 to -1.01	<0.001	3,174
Depression	0-21	-1.51	-1.63 to -1.38	<0.001	1,430	-1.92	-2.03 to -1.82	<0.001	3,174
Sleep scores	0 -100	0.16	0.14 to 0.18	<0.001	1,437	0.16	0.14 to 0.17	<0.001	3,166
Comorbidity	0-7	-1.89	-2.37 to -1.41	<0.001	1,437	-2.07	-2.47 to -1.67	<0.001	3,193
ß– Regression coefficient, CI – 95% Confidence Interval									

Table 6.16 shows multivariable regression analysis describing the relationship between baseline covariates and year 1 fatigue. Bodily pain, depression, sleep score, sex, BMI, comorbidity, and SF12Vitality at baseline were collectively associated with SF12Vitality in year one in the KPIC population and explained about half the variability of SF12Vitality (R^2=0.51) in year one. Following Vitality at baseline (std ß = 0.48), depression contributed the highest proportion to the model (std ß = -0.15), while comorbidity and BMI contributed the least to the model (std ß = -0.04 & -0.03 respectively). Anxiety at baseline was not significantly associated with SF12Vitality in year one. Linear regression assumptions were not violated, appendix figure 11.12, and multicollinearity was not detected, mean VIF=1.60.

Table 6.16: Baseline characteristics associated with year one Vitality in the total KPIC population.

	n = 4,355	**R^2 = 0.51**			
Characteristics	**Unit**	**ß**	**CI**	**p value**	**Std ß**
Vitality(b)	0 – 100	0.49	0.45 to 0.52	<0.001	0.48
Pain	0 – 100	0.07	0.04 to 0.10	<0.001	0.06
Depression	0 - 21	-0.49	-0.60 to -0.38	<0.001	-0.15
Sleep	0 – 100	0.04	0.03 to 0.06	<0.001	0.11
BMI	Kg/m^2	-0.08	-0.13 to - 0.03	<0.001	-0.03
Comorbidities	0 – 10	-0.30	-0.54 to -0.06	0.015	-0.04
Female		-1.22	-1.69 to -0.74	<0.001	-0.05

ß – Regression coefficient, Std ß – Standardized regression coefficient. Vitality(b)=Baseline vitality

Table 6.17: Multivariable linear regression model showing characteristics associated with SF12Vitality in year one, in knee pain and no-knee pain sub-populations.

		Knee Pain (n = 1,336) R^2=0.48		**No-knee pain (n = 3,019) R^2 = 0.49**	
Characteristics	Unit	ß (Confidence Interval)	Std ß	ß (Confidence Interval)	Std ß
Vitality(b)	0 - 100	0.45 (0.39 to 0.51)	0.46	0.5 (0.46 to 0.54)	0.48
Pain	0 - 100	0.08 (0.02 to 0.13)	0.07	0.07 (0.03 to 0.11)	0.06
Depression	0 - 21	-0.49 (-0.66 to -0.32)	-1.70	-0.53 (-0.67 to -0.39)	-0.15
Sleep	0 - 100	0.05 (0.03 to 0.06)	0.12	0.04 (0.03 to 0.05)	0.10
Female		-1.10 (-1.98 to -0.22)	-0.05	-1.36 (-1.93 to -0.80)	-0.06
BMI	Kg/m^2			-0.14 (-0.20 to -0.07)	-0.06
Comorbidities	0 - 10			-0.53 (-0.84 to -0.21)	-0.05

All $p \leq 0.05$, ß=Regression coefficient, std = Standardized. Vitality(b)=Baseline vitality

In the knee pain subpopulation, SF12Vitality at baseline and other baseline covariates, including bodily pain, depression, sleep score, and female sex explained about 48% of the variability of Vitality in year one (R^2 = 0.48). BMI and comorbidities were not significantly associated with SF12Vitality in the knee pain sub-population in year one. Characteristics associated with SF12Vitality in the no-knee pain sub-populations were comparable to the overall KPIC population, with bodily pain, depression, sleep score, female sex, BMI, comorbidities and SF12Vitality at baseline explaining 49% of the variability of SF12Vitality (R^2 = 0.49). In both sub-populations, following baseline SF12Vitality (std β = 0.46 & 0.48 in the knee and no-knee pain groups, respectively), depression contributed the greatest to the model (std β = -1.7 and -0.15 in the knee and no-knee pain groups respectively), table 6.17.

6.5.5 Fatigue and central sensitization

Table 6.18 represents an overview of the KPIC study population invited for further assessments (n=420). The knee pain and no-knee pain groups had similar demographic characteristics, as expected from the matching procedure. Fatigue was more prevalent in the knee pain group (77.85% vs 56.12% in the knee and no-knee pain groups, respectively) than in the no-knee pain group. SF12Vitality scores were also lower in the knee pain group, signifying more severe fatigue (40.09, SD:10.52 vs 45.99 SD, 11.21 in the knee pain and no-knee pain groups, respectively). As expected, bodily pain scores in the knee pain group were lower than in the no-knee pain group, signifying greater pain severity in the knee pain group (36.25, SD:10.00 vs 50.52 IQR: 8.37 in the knee pain and no-knee pain groups, respectively)

Table 6.18: Overview of KPIC subpopulation who had additional assessments.

Characteristics	**Unit**	**Knee pain (n=322)**	**No-Knee pain (n=98)**
Age	years	60.04 (9.63)	61.10 (9.81)
BMI	kg/m2	28.50 (25.53 to 32.99) ~	26.37 (23.87 to 29.07) ~
Sex	Female	197 (61.18) *	59 (60.20) *
Fatigue	Yes	251 (77.95) *	55 (56.12) *
Vitality	0 -100	40.09 (10.52)	45.99 (11.21)
Pain	0-100	36.25 (10.00)	50.52 (41.83 to 50.52) ~
Anxiety	0 -21	6 (4 to 10) ~	5 (2 to 8) ~
Depression	0-21	5 (3 to 8) ~	2 (1 to 4) ~
Sleep Adequacy	0-100	50.24 (28.95)	57.68 (28.41)
Ant tibia PPT	Kpa/s	397 (258 to 532) ~	377.1 (259.8 to 528.6)/404.95 (263.3 to 543.1) ~
Sternum PPT	Kpa/s	280.8 (204.8 to 391) ~	323.15 (228.35 to 401.55) ~
Index knee	right	167(52.85) *	-

* n (%), ~ -Median (IQR), knee pain and no-knee pain values are mean (SD) unless otherwise stated. SD - standard deviation, IQR- Interquartile range, PPT- Pressure Pain Detection Threshold (Distal to the most painful knee in the knee pain group & lowest value in the no- knee pain group) PPT values for no-knee pain = Left/Right

People with established knee pain, table 6.19, were comparable to the early knee pain subgroup in terms of demographic characteristics. However, the established knee pain subgroup had slightly worse patient-reported outcome measures and lower PPT values on both sites.

Table 6.19: Overview of baseline characteristics of early and established knee pain groups

Characteristics	Unit/range	Early knee pain (n=219)	Established Knee pain (n=103)
Age	years	60.00 (9.45)	59.53 (10.05)
BMI	Kg/m2	27.85 (24.85 to 31.61) ~	29.74 (26.74 to 34.83) ~
Sex	Female	134 (61.19)*	63 (61.17)*
Fatigue	Yes	159 (72.60) *	92 (89.32) *
Vitality	0 -100	41.82 (10.18)	36.40 (10.33)
Bodily Pain	0-100	39.33 (8.23)	29.64 (9.17)
Anxiety	0 -21	6 (3 to 9) ~	8 (5 to 12) ~
Depression	0- 21	4 (2 to 7) ~	6.5 (4 to 10) ~
Sleep	0-100	54.43 (28.86)	41.36 (2.19)
Ant tibial PPT	Kpa/s	399.7 (279 to 536.8) ~	315.75 (218.6 to 487.35) ~
Sternum PPT	Kpa/s	282.55 (216.3 to 396.4) ~	277.9 (166.25 to 358.15) ~
CRP	mg/l	2.05 (1.05 to 4.45) ~	2.36 (1.16 to 7.09) ~
Physical Activity	Inactive	45 (20.55) *	22 (21.36) *
	Moderately inactive	52 (23.74) *	22 (21.36) *
	Moderately active	38 (17.35) *	14 (13.59) *
	Active	80 (36.53) *	44 (42.72) *
Comorbidity	count (0-10)	1(0 - 2) ~	1 (0 - 2) ~

*- n (%), ~ - Median (IQR), Values for early and established knee pain are mean (SD) unless stated otherwise. CRP – C reactive Protein. IQR – Interquartile Range, SD – Standard Deviation

Tables 6.20 & 6.21 show that SF12Vitality was significantly correlated with BMI ($r = -0.27$, $p \leq 0.05$), anxiety ($r = -0.49$, $p \leq 0.05$), depression ($r = -0.63$, $p \leq 0.05$), CRP ($r = -0.24$, $p \leq 0.05$), and sleep ($r = 0.45$, $p \leq 0.05$). SF12Vitality was significantly weakly correlated with PPT sternum ($r = 0.21$, $p \leq 0.05$), and the anterior tibia PPT distal to the most painful knee was significantly weakly correlated with bodily pain ($r = 0.21$, $p \leq 0.05$) in the total knee pain sub-population only. None of the listed covariates demonstrated a statistically significant correlation with PPT in the no-knee pain group; however, the statistically significant high positive correlation between PPT values collected from the different sites in the knee and no-knee pain groups is noteworthy.

In the early and established knee pain subgroups (tables 6.22 & 6.23), a statistically significant correlation between PPT at both sites and SF12Vitality was not demonstrated. Furthermore, SF12Vitality was not significantly correlated with CRP, age, and BMI in both the early and established knee pain sub-populations.

Table 6.20: Table showing the paired correlation matrix for the association between SF12Vitality, PPT, and other variables in the sub-population with knee pain.

knee pain (n=252)										
	Age	BMI	Anxiety	Depression	Pain	CRP	Sleep	PPT Ant Tib	PPT Sternum	SF12Vitality
Age	1									
BMI	-0.2164*	1								
Anxiety	-0.1528	0.1824	1							
Depression	-0.1264	0.2789*	0.6031*	1						
Pain	0.0546	-0.2577*	-0.4300*	-0.5093*	1					
CRP	-0.0539	0.4733*	0.1269	0.1711	-0.2361*	1				
sleep	0.1883	-0.1900	-0.4111*	-0.4604*	0.4454*	-0.0848	1			
PPT Ant Tib	0.0227	-0.1981	-0.1419	-0.0887	0.2066*	-0.1274	0.1327	1		
PPT Sternum	0.1261	-0.2430	-0.1650	-0.1092	0.2382	-0.1578	0.1099	0.7300*	1	
SF12Vitality	0.0568	-0.2720*	-0.4852*	-0.6323*	0.4774*	-0.2430*	0.4542*	0.1551	0.2088*	1
P ≤0.05, Values are Spearman's rho(r), PPT – pressure pain detection threshold, Ant tib – Anterior tibia distal to the most painful knee, CRP – C Reactive protein.										

Table 6.21: Table showing the paired correlation matrix for the association between Vitality, PPT and other variables in the sub-population without knee pain.

No-knee pain (n=74)											
	Age	BMI	Anxiety	Depression	Pain	CRP	Sleep	PPT Ant tib lt	PPT Ant tib rt	PPT Sternum	Vitality
Age	1										
BMI	0.0806	1									
Anxiety	-0.1547	-0.1202	1								
Depression	-0.0971	0.1577	0.4218*	1							
Pain	-0.0785	-0.2234	-0.2553	-0.3798*	1						
CRP	-0.0705	0.3963*	0.012	0.2249	-0.1735	1					
Sleep	0.0146	-0.0781	-0.4012*	-0.2739	0.4148*	-0.2359	1				
PPT Ant tib lt	-0.0671	0.0927	-0.2119	-0.1577	0.1523	-0.0590	0.0462	1			
PPT Ant tib rt	-0.0596	0.1105	-0.2820	-0.2115	0.1445	-0.0735	0.1227	0.8784*	1		
PPT Sternum	0.0516	0.1230	-0.2085	-0.1417	0.0871	-0.0464	0.1843	0.7212*	0.7194*	1	
SF12Vitality	0.0656	-0.0948	-0.3392	-0.5232*	0.5282*	-0.1275	0.4676*	0.1500	0.1481	0.1843	1

P ≤0.05, Values are Spearman's rho(r), PPT – pressure pain detection threshold, Ant tib – Anterior tibia distal to the most painful knee. Lt – Left, Rt – right, CRP – C reactive protein.

Table 6.22: Correlation between vitality and covariates in the KPIC population with early knee pain

Early knee pain, n=188										
	Age	BMI	Anxiety	Depression	Pain	CRP	Sleep	PPT Ant tibia	PPT Sternum	Vitality
Age	1									
BMI	0.1509	1								
Anxiety	-0.1415	0.1031	1							
Depression	-0.0793	0.2099	0.5480*	1						
Pain	0.0445	-0.1942	-0.3025*	-0.4098*	1					
CRP	0.0497	0.4332*	0.0487	0.0997	-0.1486	1				
sleep	0.2184	-0.0859	-0.3587*	-0.4493*	0.3785*	-0.0113	1			
PPT Ant Tib	-0.0547	-0.1286	-0.1077	-0.0365	0.1415	-0.0638	0.1465	1		
PPT Sternum	0.0902	-0.1899	-0.0984	-0.0367	0.1703	-0.1393	0.1042	0.7078*	1	
SF12Vitality	0.0790	-0.2315	-0.4127*	-0.5962*	0.3746*	-0.2078	0.4380*	0.1264	0.1592	1

P ≤0.05, Values are Spearman's rho(r), PPT – pressure pain detection threshold, Ant tib – Anterior tibia distal to the most painful knee, CRP – C reactive protein.

Table 6.23: Correlation between SF12Vitality and covariates in the KPIC population with established knee pain

Established knee pain, n=64										
	Age	BMI	Anxiety	Depression	Pain	CRP	Sleep	PPT Ant Tib	PPT Sternum	Vitality
Age	1									
BMI	-0.3789	1								
Anxiety	-0.1907	0.2358	1							
Depression	-0.2535	0.3291	0.6185*	1						
Pain	0.0882	-0.3103	-0.5698*	-0.4506*	1					
CRP	-0.3466	0.5304*	0.2754	0.3067	-0.4040*	1				
sleep	0.0866	-0.3768	-0.4619*	-0.3847	0.4632*	-0.2338	1			
PPT Ant Tib	0.2608	-0.2986	-0.1151	-0.0740	0.2178	-0.2096	-0.0374	1.0000		
PPT Sternum	0.2411	-0.3419	-0.2606	-0.2024	0.3470	-0.1920	0.0179	0.7779*	1.0000	
SF12Vitality	-0.0588	-0.2739	-0.5469*	-0.6314*	0.5422*	-0.2819	0.3348	0.1333	0.3025	1.0000

P ≤0.05, Values are Spearman's rho(r), PPT – pressure pain detection threshold, Ant tib – Anterior tibia distal to the most painful knee, CRP – C reactive protein.

PPT elicited from the anterior tibia distal to the most painful knee, and sternum were significantly associated with SF12Vitality in the bivariate analysis of the total knee pain sub-population; however, after adjusting for depression and sleep, statistical significance was lost, table 6.24. Sleep, depression and PPT explained 43% and 45% of the variability of Vitality (R^2 = 0.43 in the anterior tibia and R^2= 0.45 in the sternum, respectively). This bivariate statistically significant relationship between PPT and SF12Vitality was not elicited when the population was subdivided based on the knee pain duration (early vs established) or in the no-knee pain groups, table 6.25 & 6.26.

PPT elicited from the sternum also showed a lack of a statistically significant association with SF12Vitality in the early knee pain group; however, in the established knee pain groups, statistical significance was lost in the presence of depression, table 6.26.

Table 6.24: Univariate and multivariable regression examining the association between PPT and SF12Vitality at baseline

		Total knee pain population (n=322)				
		Unit/range	Bivariate		Multivariable Model	
			ß (CI)	P value	ß (CI)	P value
*PPT Ant tibia	~PPT	kps	0.79 (0.20 to 1.37)	0.009	0.34 (-0.11 to 079)	0.181
	Depression	0 - 21			-1.48 (-1.76 to -1.20)	<0.001
	sleep	0 -100			0.74 (0.37 to 1.11)	<0.001
#PPT sternum	~PPT	kps	1.17 (0.40 to 1.93)	0.003	0.54 (-0.05 to 1.13)	0.072
	Depression	0 - 21			-1.47 (-1.75 to -1.20)	<0.001
	^sleep	0 -100			0.08 (0.04 to 0.11)	<0.001

ß - ß coefficient, CI -Confidence Interval, ~ Per 100 unit change in PPT, ^10 unit change in sleep score (*R^2 = 0.43, #R^2 = 0.45)

Table 6.25: Association between PPT and Vitality in the no-knee pain sub-population.

Subpopulation	*PPT	Unit	ß	CI	P-value
No-knee pain (n=77)	PPT Ant tib rt	kps	0.59	-0.64 to 1.82	0.341
	PPT Ant tib lt	KPs	0.77	-0.37 to 1.91	0.182
No-knee pain (n =78)	PPT sternum	Kps	0.61	-0.25 to 1.48	0.161

ß - ß coefficient, PPT measured in kps, Ant tib – Anterior tibia, lt – left, rt – right, CI -95% Confidence Interval, *Per 100 unit change in PPT.

Table 6.26: Associations between PPT and Vitality in the early and established knee pain subpopulations.

Subpopulation	PPT Site	~ß (CI)	P-value
Early knee pain (n=203)	PPT Ant tibia	0.46 (-0.20 to 1.11)	0.168
	PPT Sternum	0.71 (-0.17 to 1.59)	0.115
Established knee pain (n=72)	PPT Ant tibia	1.13 (-0.10 to 2.37)	0.071
	PPT Sternum*	0.84 (-0.33 to 2.02)	0.157

PPT measured in kps, ß - ß coefficient, CI -Confidence Interval * - After adjusting for depression only, ~Per 100 unit change in PPT.

6.5.6 Secondary Analysis

Secondary analysis using a binary (y/n) fatigue variable revealed very similar results, except that sleep was the dominant covariate in the model, table 6.27 - 6.37 and appendix figures 6.2 - 6.5. A significant relationship between PPT and fatigue was not elicited in the secondary analysis, table 6.38.

Table 6.27: Prevalence of fatigue in the total population and in people with knee pain

	Crude Prevalence	*Standardized Prevalence (CI)
Total KPIC population	62	67 (54 to 87)
Knee pain	78	84 (62 to 100)
No-knee pain	60	60 (56 to 40)

Values are %, *Standardized to Eurostat 2013 by age and sex.

Fatigue was present in over two-thirds of the total population and over 75% of the sub-population who reported knee pain, based on the y/n fatigue variable created from the KPIC dataset. These prevalences were generally higher than the prevalences derived from SF12VT in the same population; however, the prevalences in the knee pain groups were similar.

Table 6.28:Overview of people with or without fatigue in the total population

Characteristics	unit	Fatigue n=5,994	No Fatigue n=3,437
Age - M(SD)	years	61.36 (10.81)	63.28 (10.01)
BMI	kg/m2	26.95 (24.10 to 30.56) ~	25.68 (23.33 to 28.48) ~
Sex	Female	3,606 (61.07) *	1,696 (49.55) *
Vitality - M(SD)	0 -100	39.17 (11.00)	50.73 (8.20)
Pain - M(SD)	0-100	41.83 (33.14 to 50.52) ~	50.52 (41.83 - 50.52) ~
Anxiety*	0 -21	7 (4 to 10)~	4 (2 to 6) ~
Depression*	0-21	5 (2 to 8) ~	2 (1 to 3) ~
Sleep	0-100	47.09 (28.60)	75.43 (21.42)
Comorbidity count*	(0-10)	1 (0 to 2) ~	1 (0 to 1) ~
CFS		118 (1.99)	11(0.32)

*-n (%), ~ - Median (IQR), fatigue and no-fatigue values are mean (SD) unless otherwise stated. SD - standard deviation. BMI – Body Mass Index, HADs – Hospital Anxiety and Depression Scale, CFS - chronic fatigue syndrome.

Individuals with fatigue were predominantly female (68%), while the no-fatigue group had about an equal proportion of both sexes. As expected, SF12Vitality values were over 10 points lower in the fatigue group (39, SD:11) group compared with the no-fatigue group. SF12 bodily pain scores were also lower in the fatigue group (42, IQR:33 to 51).

The average bodily pain and SF12Vitality scores reported by the no-fatigue group were similar to the average UK population scores (about 50, respectively). HADs anxiety and depression scores were higher in the fatigue group compared with the no-fatigue group, with median anxiety scores of 7 IQR:4 to 10 vs 4 IQR:2 to 6 in the fatigue and no-fatigue groups, respectively, and median depression scores of 5 IQR:2 to 8 vs 2 IQR:1 to 3, in the fatigue and no-fatigue groups respectively. The fatigue group also had markedly lower mean sleep adequacy scores (47 SD:28) than the no- fatigue group, 75 SD:21. The distribution of age and body mass index was similar across the two groups.

Bivariate logistic regression

Table 6.29: Bivariate associations between fatigue and covariates at baseline in the total KPIC population

Characteristics	unit	OR	CI	P-value	n
Age	years	0.98	0.98 to 0.99	<0.001	9,323
BMI	kg/m2	1.07	1.06 to 1.08	<0.001	8,951
Sex	Female	1.60	1.47 to 1.74	<0.001	9,328
Pain	0 -100	0.38	0.36 to 0.40*	<0.001	9,283
Anxiety	0 -21	1.30	1.29 to 1.33	<0.001	9,320
Depression	0-21	1.44	1.41 to 1.47	<0.001	9,324
Sleep	0 -100	0.65	0.64 to 0.67*	<0.001	9,350
Comorbidity	count (0-7)	1.45	1.38 to 1.51	<0.001	9,381

OR – Odds ratio, CI – 95% Confidence Interval, *ORs for every 10-unit increase.

Table 6.29 shows the results of the bivariate regression analysis. Age, BMI, bodily pain, knee pain, anxiety, depression, sleep, and comorbidities were significantly associated with fatigue in the bivariate regression analysis. Table 6.30 shows baseline factors associated with fatigue resulting from multivariable logistic regression. Female sex, bodily pain, anxiety and depression, sleep, and comorbidities were significantly associated with fatigue. Table 6.31 shows the covariates associated with fatigue in ranked order according to the contribution to the model. Sleep contributed the highest to the model, while age contributed the least.

Multivariable logistic regression

Table 6.30: Factors associated with fatigue at baseline in the total KPIC population.

n=8,763				
Characteristics	unit	Total Population		
		OR	CI	p value
Age	years	0.99	0.98 to 0.99	<0.001
BMI	kg/m2	1.02	1.01 to 1.03	0.001
Sex	Female	1.32	1.23 to 1.46	<0.001
Pain	0-100	0.65	0.60 to 0.70*	<0.001
Anxiety	0 -21	1.11	1.09 to 1.12	<0.001
Depression	0-21	1.18	1.15 to 1.22	<0.001
Sleep	0-100	0.74	0.72 to 0.76*	<0.001
Comorbidity	0 - 7	1.21	1.14 to 1.32	<0.001
OR – Odds ratio, *ORs for every 10-unit increase.				

Table 6.31: Dominance analysis showing factors associated with fatigue at baseline in ranked order.

Characteristics	General dominance	Conditional dominance	Ranking
Sleep	0.1053	0.3531	1
Depression	0.0663	0.2223	2
Anxiety	0.054	0.1811	3
Pain	0.048	0.161	4
Comorbidity	0.0093	0.0311	5
BMI	0.0059	0.0197	6
Sex	0.0048	0.0161	7
Age	0.0046	0.0156	8

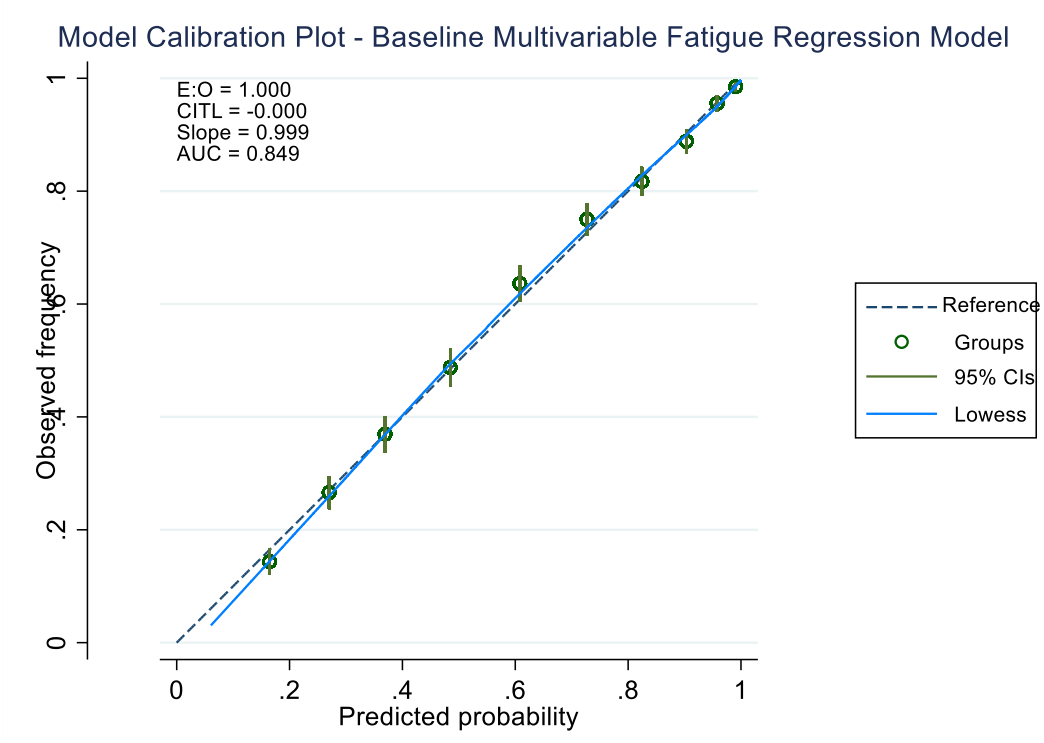


Figure 6.2:Model calibration plot

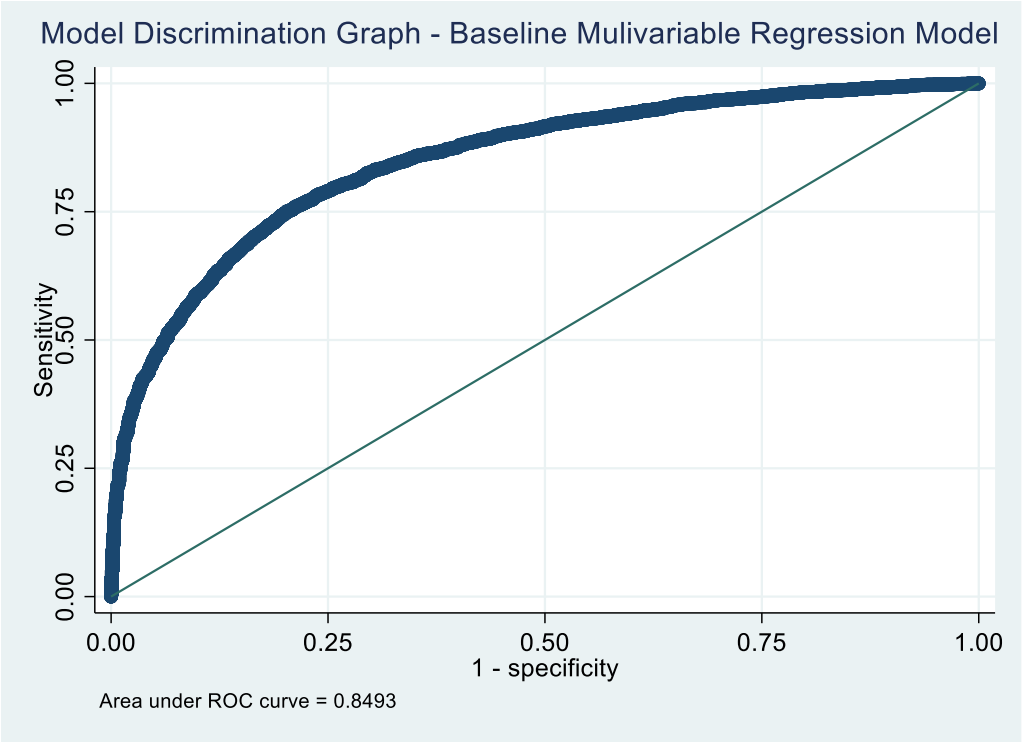


Figure 6.3:Area under the ROC curve

In Figures 6.3 and 6.4, calibration and discrimination plots for the final baseline multivariable model show good model calibration and discrimination, with calibration in the large statistic (CITL) and calibration in the slope (Slope) values of -0.000 and 0.999, and an AUROC of 0.85.

Table 6.32: Factors associated with fatigue (y/n) at baseline in the knee and no-knee pain KPIC subpopulations.

(N = 9,513)		Knee Pain (n= 2,577)		No-knee pain (n=6,317)	
Characteristics	Unit	OR (CI)	P value	OR (CI)	P value
Depression	0 - 21	1.11 (1.03 to 1.20)	<0.001	1.19 (1.17 to 1.23)	<0.001
Pain	0 – 100	0.68 (0.61 to 0.76)*	<0.001	0.66 (0.61 to 0.72)*	<0.001
Sleep	0 - 100	0.75 (0.73 to 0.78)*	<0.001	0.73 (0.71 to 0.76)*	<0.001
BMI	Kg/M^2			1.03 (1.01 to 1.04)	<0.001
Female		1.20 (0.96 to 1.51)	0.105	-1.06 (-1.46 to 1.11)	<0.001
Anxiety	0 - 21	1.16 (1.12 to 1.20)	<0.001	1.10 (1.09 to -0.05)	<0.001
Comorbidities	0 - 10			1.23 (1.12 to 1.35)	<0.001
Age	yrs.			0.99 (0.98 to 0.99)	<0.001
OR = odds ratio, CI (confidence interval) *= OR for 10-unit increase, y - yes					

Table 6.33: Table of dominance ranking (knee pain)

Characteristics	General dominance	Conditional dominance	Ranking
Sleep	0.0990	0.3591	1
Anxiety	0.0720	0.2614	2
Depression	0.0585	0.2121	3
Pain	0.048	0.161	4
Sex	0.054	0.1811	5

Female sex, sleep, anxiety, depression, comorbidities, age, BMI, and pain were associated with fatigue in the no-knee pain population. However, unlike in the no-knee pain sub-population, female sex, comorbidities, and BMI were not significantly associated with fatigue in the knee pain sub-population, table 6.32. Sleep was the dominant contributor to both final models in the knee and no-knee pain groups, followed by anxiety, table 6.33 and 6.34.

Table 6.34: Table of dominance ranking (no-knee pain)

Characteristics	General dominance	Conditional dominance	Ranking
Sleep	0.1075	0.3940	1
Anxiety	0.0608	0.2230	2
Depression	0.0467	0.1713	3
Pain	0.0359	0.1317	4
Comorbidities	0.0069	0.0255	5
Sex	0.0055	0.0201	6
Age	0.0047	0.0173	7
BMI	0.0047	0.0171	8

In year one, younger females with a higher BMI, higher anxiety and depression scores, higher comorbidities, and worse pain and sleep scores at baseline were more likely to report fatigue. Tables 6.35 and 6.36. Sleep and depression at baseline were the most significant contributors to fatigue in year one, table 6.37.

Table 6.35:Bivariate logistic describing the association between baseline covariates and fatigue in year 1 in the total KPIC population.

Characteristics	unit	OR (CI)	P-value	n
Age	years	0.99 (0.98 to 0.99)	<0.001	4,631`
BMI	kg/m2	1.07 (1.06 to 1.09)	<0.001	4,443
Sex	Female	1.60 (1.47 to 1.74)	<0.001	9,328
Pain	0-100	0.38 (0.35 to 0.42)*	<0.001	4,600
Anxiety	0 -21	1.26 (1.24 to 1.29)	<0.001	4,610
Depression	0-21	1.38 (1.35 to 1.42)	<0.001	4,610
Sleep	0 -100	0.71 (0.69 to 0.73)*	<0.001	4,608
Comorbidity count	0-10	1.47 (1.38 to 1.57)	<0.001	4,634

OR – Odds ratio, CI – 95% Confidence Interval, *ORs for every 10-unit increase.

Table 6.36:Baseline covariates associated with fatigue in year 1 in the multivariable logistic regression model in the total KPIC population.

n=4,531 Year 1 fatigue				
Characteristics	Unit	OR	CI	P value
Age	0 - 100	0.99	0.98 to 1.00	0.003
Sex	0 -21	1.33	1.21 to 1.46	<0.001
Pain	0-100	0.60	0.54 to 0.67*	<0.001
Anxiety	0 - 21	1.09	1.05 to 1.14	<0.001
Depression	0 -21	1.15	1.11 to 1.18	<0.001
Sleep	0 - 100	0.81	0.80 to 0.82*	<0.001
Comorbidity	0 – 10	1.25	1.12 to 1.41	<0.001
all $p < 0.005$, OR - Odds Ratio, CI - 95% Confidence interval, *ORs for every 10-unit increase.				

Table 6.37:Dominance analysis showing baseline factors associated with fatigue in ranked order.

Characteristics	General dominance	Conditional dominance	Ranking
Sleep	0.0659	0.284	1
Depression	0.0529	0.2278	2
Pain	0.0509	0.2195	3
Anxiety	0.0427	0.184	4
Comorbidity	0.012	0.0516	5
Age	0.003	0.0128	7
Sex	0.0047	0.0203	6

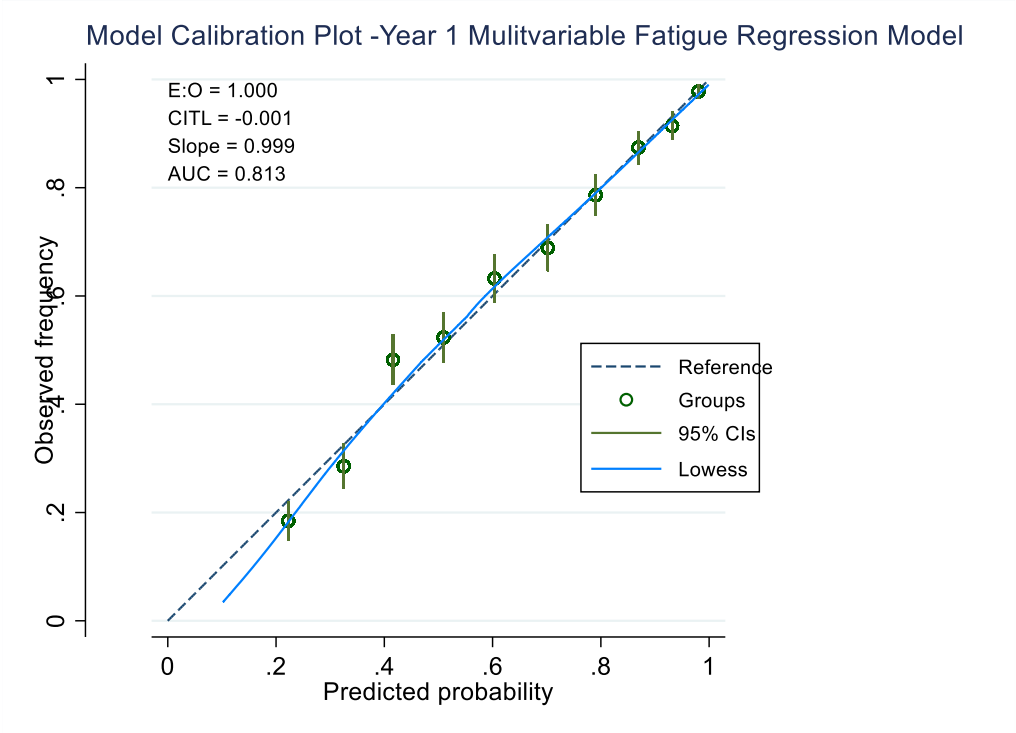


Figure 6.4:Model calibration Plot

Figure 6.5 shows the calibration plot for the multivariable model describing baseline covariates associated with fatigue in year one. It shows good model calibration and discrimination with CITL and slope values of -0.001 and 0.999 respectively. AUROC was 0.81, figure 6.6.

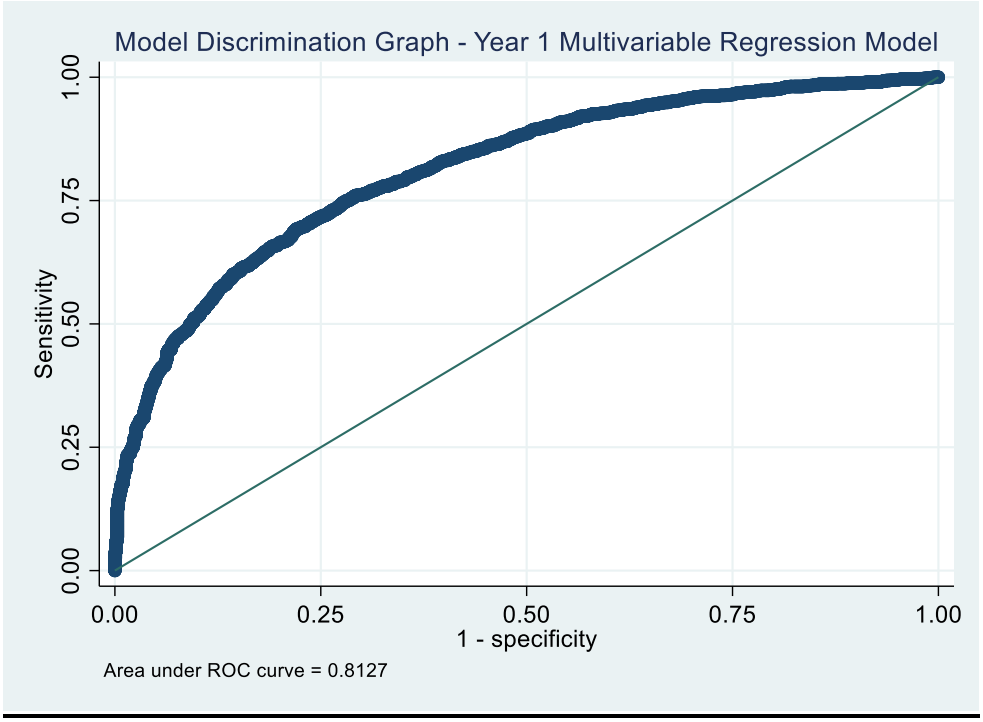


Figure 6.5:Area under the ROC curve

Association between fatigue and central sensitisation

Table 6.38 shows that a statistically significant relationship was not elicited between fatigue and Anterior Tibia PPT in the knee pain and no-knee pain groups, and in people with early and established knee pain. PPT sternum was however associated with fatigue in the univariate analysis but not in multivariable after adjusting for sleep in the no-knee pain group.

	Sub-population		OR (CI)	P-value	n
PPT Ant tibia	Total knee pain		0.92 (0.81 to 1.04)	0.191	275
	Early knee pain		0.96 (0.83 to 1.10)	0.528	203
	Established knee pain		0.85 (0.61 to 1.19)	0.344	72
	No-Knee pain	lt	0.80 (0.64 to 1.00)	0.050	77
		rt	0.79 (0.63 to 1.00)	0.052	77
PPT Sternum	Total knee pain		0.90 (0.76 to 1.06)	0.186	278
	Early knee pain		0.93 (0.78 to 1.12)	0.476	206
	Established knee pain		0.97 (0.81 to 1.17)	0.226	72
	*No-knee pain		0.71(0.46 to 1.20)	0.123	78

PPT – Pressure Pain Detection Threshold, CI – Confidence Interval, lt – Left, rt-Right, OR – Odds Ratio for 100 unit change in PPT, *after adjusting for sleep

Table 6.38: Association between fatigue and PPT in KPIC subpopulation who had further assessments.

6.6 Discussion

6.6.1 Key findings and interpretations

This chapter examined fatigue in the general population over 40 years using data from a cohort of adults in the East Midlands. Findings revealed that fatigue is prevalent in the general adult population over 40 years, and patient-reported outcome measures of pain, sleep, depression, and female sex were consistently associated with fatigue and predicted persistent fatigue in the general population. Interestingly, the association between pressure pain detection thresholds in the two sites assessed (anterior tibia and sternum) were not consistently associated with fatigue.

Prior studies have demonstrated the association between female sex and fatigue in the general population(Galland-Decker et al., 2019, Lerdal et al., 2005). This study also accords with earlier observations, which demonstrated the association between fatigue and female sex, pain, and mental health in people with rheumatoid arthritis.

The symptoms of pain, depression, and poor sleep have been noted to co-occur in many disease states associated with fatigue, more frequently than can be explained by chance alone (Gaston-Johansson et al., 1999, Reyes-Gibby et al., 2006, Forbes et al., 2006, Laird et al., 2011). In this study, I demonstrated a similar symptom cluster in the general population, supporting my hypothesis that these symptoms can occur in the absence of a specific disease state. Aberrant production of proinflammatory cytokines has been aetiologically linked to this cluster of symptoms; however, the symptom cluster has been described in non-

inflammatory conditions and in our study of a general population sample with a low prevalence of the inflammatory disease. This makes a primary inflammatory cause less likely, and in addition, perhaps not surprisingly, an association between fatigue and a measure of inflammation was not elicited in our study. I, therefore, suggest that a specific phenotype may be associated with this symptom cluster and more likely to develop fatigue with inflammatory or non-inflammatory diseases.

As highlighted in earlier sections of this chapter, abnormal central processing is associated with the afore mentioned symptom cluster and could be implicated in the pathophysiology of fatigue in many disease conditions. Central sensitisation could be the common thread that defines this phenotype as more susceptible to developing persistent fatigue. Although the association between PPT and fatigue did not consistently achieve statistical significance, more research would be required to test this theory in a larger population or in a more extensively characterised sample, for example more modalities of ppt, other measures of central sensitisation in a larger sample. This may uncover other causal mechanisms and possibly develop interventions that could mitigate the process or alleviate symptoms.

Another interesting finding of this study is the dominance of the association between depression and sleep and fatigue over and above other measures in the main and secondary analysis, respectively. The contribution of sleep and depression could be explained perhaps by shared neurophysiology through the prefrontal cortex and hypothalamus(Koenigs and Grafman, 2009, Stenberg, 2007). Alternatively,

each analysis could be describing a single dimension of fatigue, given its multidimensional nature, where different unidimensional instruments may ascribe dominance to measures associated with one dimension of fatigue.

Contrary to our expectations, however, this study did not elicit a significant association between PPT and fatigue in the population, and although a significant association in people with knee pain was ascertained, it was not sustained in multivariable analysis. A possible explanation may be that fatigue, although thought to be associated with heterogeneous brain structures, is commonly linked to the frontal and parietal lobes, limbic system, and basal ganglia, while PPT is known to activate the spinothalamic tract, responsible for pressure and pain sensation through Aδ, and C afferents (Backonja et al., 2013, Goñi et al., 2018). The lack of statistically significant associations between PPT and fatigue in people with early or established knee pain could be explained by the smaller sample sizes and a study population not powered to detect an effect of one being present. It is important to bear in mind that several measuring tools can be used to elicit the presence of central sensitisation, and the results of a single modality do not lend enough evidence to confirm or refute the findings. More studies would be required on a larger sample, using the full complement of QST and other measures of central sensitisation.

6.6.2 Strengths and Limitations

One of the strengths of this study is the large sample size of over 9000 participants and the sampling strategy, which allowed for the collection of data from people in the community. This facilitates generalisability to the target population and detection of effects if present. The study also benefits from robust statistical analysis and subgroup and secondary analyses to mitigate some potential sources of bias that could result from the sampling framework. PPT was also measured from two sites.

The study is not without limitations. Firstly, findings were obtained from a population in the East Midlands, United Kingdom and may not be generalisable to the global population. However, according to data obtained from the office of national statistics, demographic characteristics of the east midland population do not differ significantly from the UK population (Office of National Statistics, 2023). Back weighted KPIC data to the east midlands population yielded nearly identical results (Appendix 11.6.1). In addition, age and sex standardisation would allow comparison with other populations.

The effect of selective non-response deserves some consideration. Overall KPIC response rate was less than 30% at baseline. I calculated weighted estimates to match the age and sex distribution of general population of the East Midlands, age-sex weighted analysis produced similar findings to the unweighted estimates. This was probably because, several studies, including the ERAN study showed the lack of a significant association between age and fatigue. In addition, sex was

accounted for during multivariable modelling. Be that as it may, it was impossible to statistically determine the effect of non-response on the outcome variable because the distribution of other characteristics in the UK population (e.g., pain and sleep levels) were unknown. Studies however have shown that voluntary respondents of surveys indicate more favourable health outcomes e.g., better mental health and lower alcohol consumption (Van Loon et al., 2003, Criqui et al., 1978, Cheung et al., 2017). This indicates that our study findings may well be conservative.

Conversely, the name of the study could inadvertently introduce selection bias towards people with knee pain, which may over estimate measures of effect related to fatigue. I accounted for this possibility by conducting a subgroup analysis on the sample stratified by the presence and absence of knee pain. An alternative way may have been the implementation of multivariable modelling using interaction terms. Stratification and interaction terms are valid statistical methods, however, in this thesis, stratification by knee pain was preferred over a knee pain interaction term. Although interaction terms use all data, and may be more computationally efficient, stratification is simple and does not make implicit assumptions about the association between other predictors in the model. This facilitated the estimation of risk factors associated with fatigue in people with knee pain, independent of those without knee pain (Greenland, 1989, Corraini et al., 2017). Multivariable modelling would have required several interaction terms consequently reducing the statistical power of the study. Additionally, stratification is recommended

where there are few stratifying variables, such as in this analysis, where data were stratified by only one variable (Tripepi et al., 2010b).

The lack of objective measures of fatigue and the absence of clearly defined cut-off points to delineate the presence or absence of fatigue could lead to fatigue misclassification; however, a validated measure of fatigue was utilised, and another measure of fatigue with good face validity was used to triangulate findings in the secondary analysis.

6.7 Conclusions

Fatigue is prevalent in the general population. Traits related to central sensitisation are associated with fatigue and predict persistent fatigue. Depression and sleep are dominant symptoms associated with fatigue in the general population. Populations with worse patient-reported outcomes such as poor sleep, mental health and pain could represent a phenotype more susceptible to developing persistent fatigue and could represent targets for future fatigue research and intensive fatigue lowering interventions in the presence of disease. Further studies are required to evaluate the mechanisms of fatigue and the development of tools to investigate central aspects of fatigue in people with musculoskeletal diseases.

7. Fatigue and pain in rheumatoid arthritis: investigating the contribution of central mechanisms

7.1 Outline

In previous sections of this thesis, I investigated fatigue in rheumatoid arthritis and in the general population. I established the prevalence of fatigue and identified possible links between fatigue and central sensitisation in rheumatoid arthritis and in a sample of the general population. Interestingly, the association between fatigue and a measure of central sensitisation, pressure pain detection threshold (PPT), was not consistent in the general population sample. This left unanswered questions about the usefulness of measures of central sensitisation in fatigue research and the relationship between fatigue and central sensitisation in other disease conditions, and in the general population. Additionally, in previous chapters of the thesis, I utilised secondary data analysis to answer the research questions.

In this chapter, however, I described the protocol development and set-up of a cohort study that investigates the contribution of central sensitisation to pain and fatigue in RA. Building on the information uncovered so far in this thesis, this study was designed to evaluate the contribution of central sensitisation to fatigue and pain in RA.

7.2 Introduction

My previous work demonstrated that fatigue is prevalent and persistent in people with RA. I described a link between fatigue and traits associated

with abnormal central processing and central sensitisation. Individuals with these traits were more likely to experience fatigue and worse clinical outcomes in RA (Ifesemen et al., 2021) and in the general population.

Abnormal central processing of noxious stimuli is thought to be a consequence of long-standing changes to parts of the central nervous system that could be linked to the persistence of pain, depression, and fatigue in people with many chronic disease conditions. This is particularly important in inflammatory conditions such as RA, where clinical indices used to monitor disease activity often include patient-reported outcome measures such as pain(Anderson et al., 2012).

Abnormal central processing could be the missing link between fatigue and central sensitisation and explain the concurrent association between fatigue and pain in RA (Baek et al., 2020, McBeth et al., 2015). Refractory pain without the concordant signs of inflammation would lead to a spurious increase in disease activity scores, for example, DAS28 in the presence of well-controlled inflammatory disease in subgroups of individuals(McWilliams and Walsh, 2016, McWilliams et al., 2019). In these circumstances, misclassification of active disease could result in the increased use of DMARDs, exposing people to the risk of side effects without the real prospect of benefit(Lage-Hansen et al., 2016). Too narrow a focus on suppressing inflammation could displace the introduction of adjunctive treatments targeted at central sensitisation and possibly cause poorer long-term outcomes in people with RA.

Treatment of RA currently adopts a ‘one size fits all’ approach to the management of pain and fatigue, mainly targeted at inflammatory

disease and nociceptive pain(Deighton et al., 2009). Fatigue is largely ignored in the consulting rooms. However, different pain mechanisms represent targets for targeted pain management and might permit stratified or precision treatment, and specialised packages of care could be developed for people with fatigue. Furthermore, available evidence describes a heterogenous RA population with varying prognosis and treatment responses in terms of disease activity, pain, fatigue and functional limitations, and psychological distress (Wabe et al., 2017, Druce et al., 2015b, McWilliams et al., 2019, Norton et al., 2013, Pan et al., 2019, Ifesemen et al., 2021). For example, people with high inflammatory activity might benefit from a change in DMARD strategy, whereas those with pain and fatigue from other causes might benefit from other adjunctive treatments.

A stratification tool that identifies people with abnormal central processing could represent an evolution in managing people with RA and many other inflammatory diseases. Drawing parallels from pain research, where abnormal central processing has been extensively studied, it is known that abnormal central processing can be identified using a variety of methods, including quantitative sensory testing (QST) and brain imaging (Uddin and MacDermid, 2016, Georgopoulos et al., 2019, Harte et al., 2018). Though useful in research, these methods have little utility in clinical settings, for example, in general practice (GP) clinics (Uddin and MacDermid, 2016). Questionnaires permit economical assessment with relative ease and are suitable for use with large populations and in busy clinical settings. Questionnaires that can assess abnormal central

processing should, therefore, be more amenable than QST or brain imaging for standard applications in clinics and in research.

A single-page eight-item questionnaire, central aspects of pain in the knee (CAP-KNEE), was designed and validated to measure traits associated with abnormal central processing in people with knee pain. This self-report tool measures eight characteristics strongly associated with central pain mechanisms and predicts 1-year pain outcomes in people with knee pain (Akin-Akinyosoye et al., 2020) (Akin-Akinyosoye et al., 2021). The central aspect of pain in RA (CAP-RA) is a minor adaptation of the CAP-KNEE questionnaire for use in a RA population. I hypothesise that fatigue is associated with central sensitisation, and traits associated with abnormal central processing, measured by the CAP-RA questionnaire, are associated with fatigue in RA. Therefore, this questionnaire could be valuable for pain and fatigue research.

7.2.1 Aims and objectives

7.2.1.1 Aims

This study aims to further develop and validate the new self-report 8-item CAP-RA questionnaire for measuring central pain mechanisms in RA and investigate the association between central sensitisation and fatigue in RA. The working hypothesis is that the CAP-RA score is associated with fatigue in RA.

7.2.1.2 Objectives

The primary objective of this study is to optimise and establish the psychometric properties of CAP-RA.

Secondary objectives are to:

- To measure the ability of CAP-RA to identify people with fatigue and pain augmented by central mechanisms who may have poorer outcomes despite modern therapy.
- Investigate factors associated with worse RA pain or fatigue at baseline and three months follow up.
- Derive CAP-RA scoring recommendations for stratification in clinical trials and in clinical practice.
- Examine associations between pain, central sensitisation, and fatigue in RA.
- Examine the short-term course of pain and fatigue in RA.

The rest of the chapter describes the study protocol of the CAP-RA study, reported according to the SPIRIT guidelines (Chan et al., 2013), Appendix 11.5.1.

7.3 Methods

7.3.1 Study design and setting

This is set up as a prospective observational study based in secondary care. Adults with RA who meet the study inclusion criteria will be recruited through secondary care rheumatology services (Sherwood Forest Hospitals NHS Foundation Trust; SFH). Participants will undergo clinical assessments, blood tests, ultrasound scans and complete the study questionnaire booklet at baseline and twelve weeks later. A consecutive sample of participants will, in addition, complete the CAP-RA questionnaire one week after the baseline visit.

7.3.2 Ethical considerations

At the time of drafting this thesis, the study protocol had been reviewed and given a favourable ethical opinion by the North of Scotland research ethics committee (1) (reference no: 20/NS/0036) and approved by the health research authority (HRA). As described in the ethics application, the CAP-RA study will be performed in accordance with the UK policy framework for health and social care research, 2018, principles of good clinical practice and Helsinki declaration.

Informed consent will be sought and confirmed through a signed consent form at the start of the first study visit before commencing study procedures. Participants would be explicitly informed that partaking in the study would not affect the participant's entitlements, treatment, or care during the study or in the future; and would not require or preclude participation in other research projects.

7.3.3 Public and patient involvement and engagement in the CAP-RA study

The PPI component of this study was conducted prior to writing up this thesis and describes the involvement process in the study design and procedures stage of the execution phase of this research project. PPI was described according to the procedure outlined by Shippee et al. and reported according to the guidance for reporting involvement of patients and the public (GRIPPS 2) short form checklist (Staniszewska et al., 2017).

The CAP-RA research team made up of my supervisors (DMcW and DAW) and collaborators EF, KAA, DW, RW, the lay collaborator DP, and I co-developed this CAP-RA study. A layperson was a co-applicant on the study funding application.

Following this, a study steering committee comprising two lay members was constituted to oversee all aspects of the study. Terms of reference can be found in appendix 11.5.2. Lay members of the committee received brief training on the research process and requirements for each stage of a research project from a member of the research team. A plan was instituted to assess the training needs of the research team, including lay members throughout the project's life course, and requisite training was administered accordingly.

A Lay member and researchers co-produced the first iteration of the study design. A PPI event was then organised to engage with patients and lay members of the public to discuss the CAP-RA questionnaire and the overall study design. Participants were recruited through snowball sampling by DP (PPI and CAP-RA steering committee members). Participants willing to participate in the event were invited to a focus-group style meeting to discuss the CAP-RA questionnaire and overall study design. Firstly, each participant completed the CAP-RA questionnaire. A focus-group discussion followed to talk about their understanding of the questionnaire and their experience completing it. Talking points, adapted from the CAP-KNEE questionnaire development, were utilised to stimulate discussion. Where possible, open-ended

questions were utilised, and follow-up questions were asked to clarify answers.

The meeting took place at the Kings Mill Hospital in February 2020, was chaired by me and attended by six lay participants, two research nurses, two steering committee PPI members, and an experienced PPI facilitator (to monitor the discussions and act as a scribe).

Overall, participants thought the study was necessary and understood the meaning of the questions in the questionnaire (appendix 11.5.3). Following this meeting, the study design was revised, and participant information sheets, consent forms and letters to participants were developed (appendix 11.5.4). Patient partners reviewed and approved all study materials and included specific instructions; for example, one significant change was the inclusion of the following sentence -

"…. *that participation in the study would not affect the participants' entitlement,"*

This addition was made in response to feedback from PPI about putting assurances in the study information sheets that participation in the study would not affect the disability benefits some participants may receive because of their medical condition. Lay members of the study steering committee also co-authored the study protocol manuscript(Ifesemen et al., 2021).

I believe that patient public involvement validated the relevance of this study process and would add credibility to the study findings. PPI was conducted at every stage of the study design, including utilising the social

networks of the lay members of the steering committee to recruit participants for the focus group discussion. This process also served to educate the lay steering committee members on the life course of a research study. Finally, on personal reflection it provided beneficial experiential learning on the value and contribution of PPI in a research study.

Although this PPI event provided an invaluable contribution to this research study, there was room for improvement. The group would have benefited from some prior training on the conduct of focus group discussions because sometimes the participants seemed to focus on other topics outside the remit of the research study. Additionally, although the group represented the demographic distribution of people commonly affected by RA, my sampling technique meant that this group might possess similar epistemological perspectives. A more diverse group may have provided a broader perspective.

7.3.4 CAP-RA study participants

Participants will be included in the study if they were aged 18 years and above, satisfied the European League Against Rheumatism (EULAR) criteria for RA (Aletaha et al., 2010), and the presence of active RA, defined as DAS28 ≥ 3.2 at the baseline visit, pain score >3 on a 0 – 10 numerical rating scale where 0 = no pain and 10 = worst imaginable pain. Participants will be excluded from the study if they are unable to give informed consent, unable to understand sufficient spoken or written English to comply with the requirements of the study protocol, unable or unlikely to complete the proposed 12-week study follow-up (e.g. moving

house, terminal diagnosis, current or planned pregnancy), have active co-morbidity (e.g. uncontrolled diabetes mellitus, cancer, infection) requiring changes in medical treatment at baseline, major active psychiatric condition (e.g. major depression) and inability to meet the requirements of clinical assessments.

7.3.4.1 Recruitment

Participants would be recruited from secondary care through different modalities. In the first instance, outpatient records will be screened by a member of the clinical care team, and eligible participants will be contacted by letter and invited to join the study. Posters advertising the study (appendix 11.5.4.1) will be placed in public areas within clinics/hospitals, alongside researchers' contact details, allowing potential participants to contact the research team directly. Additionally, members of the research team will recruit participants via face-to-face discussions in Rheumatology outpatient clinics as permitted by local policies.

7.3.4.2 Study regimen

Participants enrolled in the study will be assessed at study visits conducted at baseline and 12 weeks later. The following section describes the CAP-RA study regimen in detail. The study assessments at baseline and 12 weeks will have identical data collection protocols. One week after the baseline assessment, a subset of 80 consecutive participants will be invited to complete and return a postal questionnaire consisting of the 8-item CAP-RA questionnaire to examine the test-retest

reliability of the instrument. Participants will be asked to opt-in to respond to weekly short message service (SMS) text messages asking about average fatigue and pain levels. The study regimen is described in Figure 7.1.

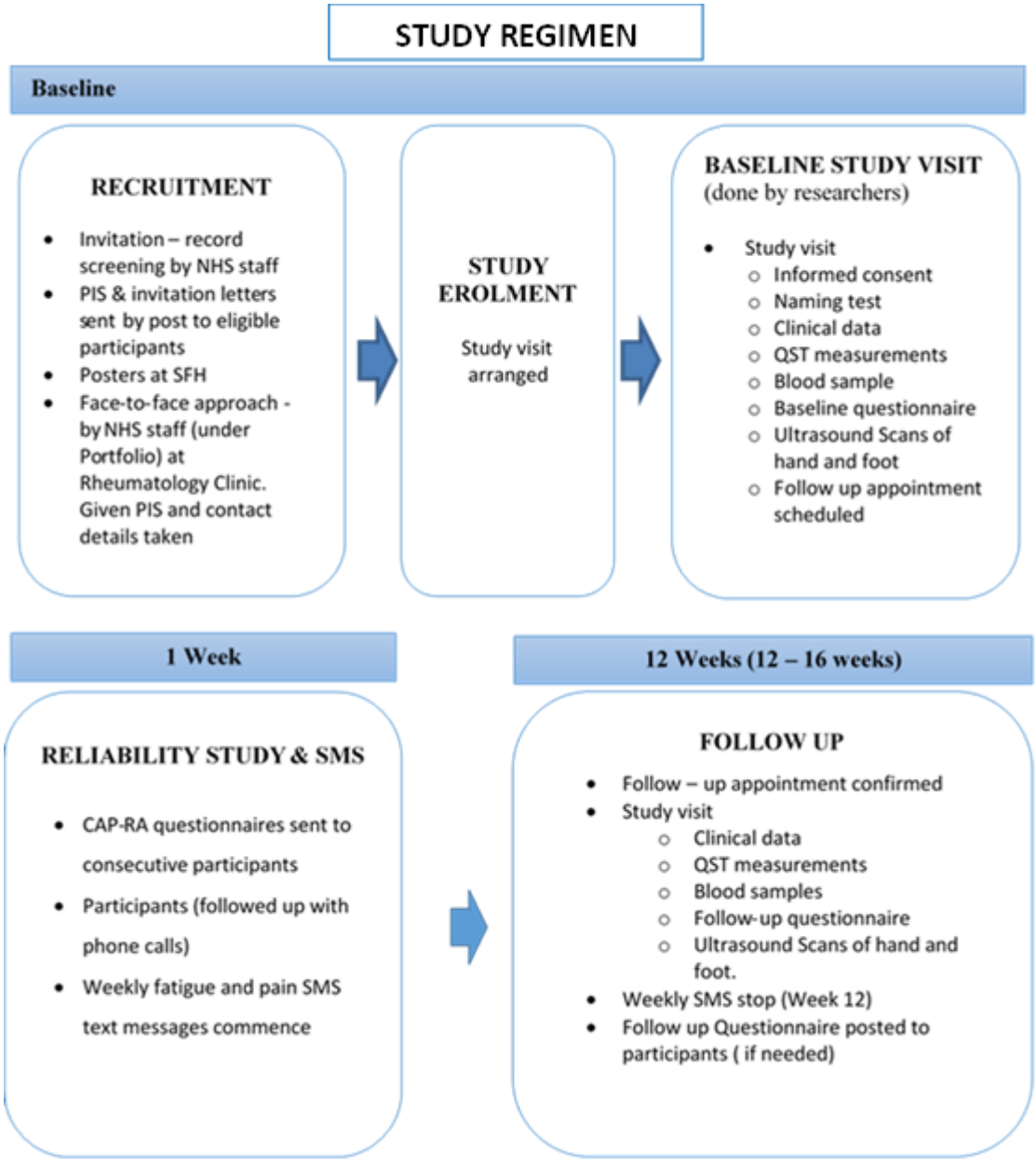


Figure 7.1: CAP-RA study regimen

7.3.5 Study data collection

Study data would be collected through questionnaires, clinical and laboratory assessments, and imaging.

7.3.5.1 Questionnaire

The study questionnaire was also developed prior to writing this thesis. It comprised a collection of validated instruments and the newly adapted CAP-RA questionnaire. The questionnaire configuration is described below:

In the questionnaire, fatigue will be measured by Bristol RA Fatigue Scale (BRAFs) multidimensional questionnaire and pain, measured using a 0 – 10 numerical rating scale (Hewlett et al., 2018).

Traits associated with central pain mechanisms will be assessed using the CAP-RA questionnaire and the central sensitisation inventory short form-9(CSI-9)(Scerbo et al., 2018). Information on the various pain mechanisms will be collected by the different validated questionnaires. Neuropathic-like pain will be collected by a modified painDETECT questionnaire(Rifbjerg-Madsen et al., 2017). Modifications to this questionnaire include the exclusion of the pain waveform and pain intensity questions because the aforementioned items are not included in the neuropathic pain score. painDETECT questions were reworded as statements for easy readability. In addition, pain catastrophising will be measured using the pain catastrophising scale (PCS) (Osman et al., 2000).

Other measures collected by the questionnaire will include depression and anxiety, measured using the hospital anxiety and depression scale (HADS)(Zigmond and Snaith, 1983). Functional ability and sleep will be measured by the Athens insomnia scale (AIC)(Soldatos et al., 2000). Cognitive function will be measured using the cognitive failures questionnaire (CFQ)(Broadbent et al., 1982).

Pain distribution from the drawing of a whole-body manikin would be used to derive the widespread pain index (WPI), which, in addition to the symptom severity score (SSS), would be used to classify the presence of fibromyalgia (Wolfe et al., 2016, Wolfe et al., 2010). Disability and physical activity will be assessed with the health assessment questionnaire (HAQ) (Maska et al., 2011) and the short international physical activity questionnaires (IPAQ) (Cleland et al., 2018), respectively. Demographic characteristics will be recorded, and comorbidities required to calculate the rheumatic disease co-morbidity index (RDCI) will be assessed (England et al., 2015). A detailed medication history will be collected, including common medications used to manage RA and common pain medications, e.g. nonsteroidal anti-inflammatory drugs (NSAIDs) and paracetamol. Appendix 11.5.4.5 contains the opening pages of the CAP-RA questionnaire booklet.

7.3.5.2 Clinical Assessment

Clinical assessments will include the RA disease activity score in 28 joints (DAS28) and three modalities of quantitative sensory testing, including pressure pain detection threshold (PPT), temporal summation (TS), and conditioned pain modulation (CPM).

Patient global assessment (PGA) of disease activity will be assessed using a visual analogue scale from 0 – 100 on an incremental scale. DAS28-Erythrocyte Sedimentation Rate (ESR) will be estimated using the published formula.

7.3.5.1.1 DAS - 28 assessment

Section 7.3.8 describes a brief quality control reliability study conducted on consecutive patients with RA in a rheumatology outpatient clinic during my first training session for TJC and SJC assessment. My trainer was my supervisor, Professor David Walsh, an experienced rheumatologist who performs these assessments routinely in the clinic.

7.3.5.1.2 Quantitative sensory testing (QST)

As mentioned earlier, QST will be conducted using pressure pain detection threshold (PPT), temporal summation (TS) and conditioned pain modulation (CPM) tests. PPT involves the assessment of mechanical pressure pain thresholds using a handheld computerised pressure algometer (Medoc Algomed – Computerised Pressure Algometer). The algometer consists of a rod with a 1cm diameter circular end and is applied perpendicular to the test site. Pressure is applied through the algometer at a standardised rate until the participant presses a button, indicating a change in sensation from pressure to pain. The algometer is immediately removed, and the pressure reading is recorded. Three measurements will be collected at each site, with a short rest between measurements. Participants will be familiarized with the procedure by performing it on a different test site before commencing the test. PPT will be conducted on three anatomical sites: the body of the

brachioradialis muscle of the non-dominant forearm, the medial tibiofemoral joint line, and the body of the anterior tibialis muscle of the dominant side (Rolke et al., 2006). PPT is measured in kilopascals (kPa), with low values indicating more pain sensitivity.

Temporal summation (TS) tests the nervous system's responsiveness to repeated noxious stimuli. It will be conducted using a supra-threshold pinprick stimulator (256mN pinprick). The pinprick simulator will be applied repeatedly to a single point on the skin over the patella ligament in a single test and on a train of 10 repetitions, with intervals of 1 second in between repetitions. The participants will be asked to indicate the intensity of pain on a 0 – 100 visual analogue scale (where 0 represents no pain and 100, the most intense pain imaginable) for the single test and the average pain intensity for the series. The difference between the two self-reports will be used to measure TS.

Conditioned Pain Modulation (CPM) is conducted using the PPT test (test stimulus). The test stimulus is applied to the body of the anterior tibialis muscle contralateral to the participant's dominant hand before and after induction of a painful stimulus at a remote anatomical site. This conditioning stimulus is achieved by applying a blood pressure cuff to the dominant arm and inflating it until the radial pulse can no longer be palpated. The participant may be asked to perform hand exercises (squeezing a stress ball) to induce exercise-induced ischaemic pain or discomfort. As soon as the pain or discomfort in the forearm is rated as four out of ten by the participant, they stop squeezing the ball, and the

second PPT test is performed. Only one CPM test will be performed. CPM is calculated as (Yarnitsky et al., 2015):

Equation 7.1: CPM formula

$$CPM = PPTpost - PPTpre$$

7.3.5.3 Imaging

Joint inflammation (synovitis and tenosynovitis) will be assessed by ultrasound scans of the hands and feet in both grayscale (GS) and power doppler (PD) modes. Examinations will be conducted according to the European league against rheumatism (EULAR) guidelines for MSK ultrasound scans in rheumatology (Backhaus et al., 2001). Joint inflammation will be assessed using the outcome measures in rheumatology (OMERACT) US group scoring criteria (Naredo et al., 2013, D'Agostino et al., 2017).

7.3.5.4 Laboratory Assessments

Blood samples will be obtained for inflammatory biomarkers (erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) assays, and samples may be stored for additional biomarker analyses.

7.3.5.5 Additional data collection

The CAP-RA questionnaire will, in addition, be mailed for completion by 80 consecutive participants one week after the baseline visit to assess test-retest reliability.

The course of fatigue and pain will be assessed by sending participants weekly SMS text messages at approximately the same time each week. The text message will contain the following question: “On a scale of 0 –

10, with 0 indicating no fatigue and 10 indicating the worst imaginable fatigue, how would you rate your fatigue during the past week? Please respond to this message with a number. Thank you". The identical question wording will be used to ask about weekly pain.

7.3.6 Statistical analysis

The primary objective of the study is to determine the psychometric properties of the new CAP-RA questionnaire. Statistical methodology used to examine the association between fatigue and central sensitisation has been described in Chapter 3. The following describes the proposed analysis methodology to validate the CAP-RA questionnaire.

Psychometric properties of CAP-RA examined in this study will include structural validity, construct hypothesis testing, internal consistency, reliability, measurement error and criterion validity. These properties would be assessed in accordance with consensus-based standards for the selection of health measurement instruments (COSMIN) guidelines (Prinsen et al., 2018, Terwee et al., 2018, Mokkink et al., 2018).

Structural validity measures the degree to which the scores in a patient-reported outcome measure reflect the dimensionality of the construct to be measured. Structural validity will be assessed using confirmatory factor analysis (CFA) and Rasch modelling. A suitable polytomous Rasch model will be selected(Mokkink et al., 2018). The unidimensional and local independence assumptions will be examined, and model fit parameters will be estimated(Tennant and Conaghan, 2007). The construct validity of

the CAP-RA questionnaire, a measure of the degree to which scores are consistent with a hypothesis that assumes that the PROM validly measures the intended construct, will be investigated using confirmatory factor analysis. Appropriate fit statistics will be calculated(Atkinson et al., 2011). Internal consistency, which is a measure of the degree of interrelatedness among items in the questionnaire, will be estimated using Cronbach's alpha, with a value greater than 0.70 signifying acceptable internal consistency(Streiner et al., 2015).

Test-retest reliability will be estimated using the intraclass correlation coefficient, with acceptable values ≥0.70(Koo and Li, 2016). Measurement error will be estimated by the standard error of measurement (SEM), smallest detectable change (SDC) and the limits of agreement (LoA). Criterion validity will be a measure of the degree to which the instrument is an adequate reflection of central sensitisation.

The relationship between baseline CAP-RA scores and pain outcomes will be measured using regression analysis, with 12-week pain as the outcome variable and baseline characteristics as covariates.

Secondary objectives will be addressed using appropriate regression analysis. For example, the association between fatigue and central sensitisation measured using PPT will be examined using linear or non-linear regression analysis based on the distribution of the data, and the course of fatigue will be examined using linear or non-linear mixed effect models as indicated. CAP-RA questionnaire thresholds for delimiting high levels of central mechanisms will be derived through receiver

operating curves plotted using dichotomised outcome measures, such as the different QST measurements (Mokkink et al., 2018, Hill et al., 2008).

Analyses will be undertaken using appropriate statistical software. P values less than or equal to 0.05 will be considered statistically significant unless otherwise stated.

7.3.7 Sample size considerations

The overall study sample size was calculated to determine the structural validity of the instrument using Rasch analysis and conducted during the design phase of this study. The Rasch model assumes that item calibrations are within ±½ logit from stable values, with a 99% confidence interval. This gave an overall sample size of between 108 and 243, best to poor targeting(JM, 1994). I expect 200 participants from a sample of 250 participants to complete the study. This is in keeping with COSMIN recommendations for Rasch analysis(Prinsen et al., 2018, Terwee et al., 2018, Mokkink et al., 2018).

For longitudinal analysis over 12 weeks, I predict that n = 200 people will complete both study visits. The power calculation for multiple linear regression with CAP-RA plus five baseline covariates (each assumed to have a correlation coefficient of 0.15 with the 12-week pain) yielded a sample size of n=164 for α of 0.05 and power of 0.95.

Seventy-seven participants are required for the reliability study to estimate an ICC of 0.8, with two ratings, alpha of 0.05, 80% power, and a lower limit ICC not less than 0.6. Sample size calculation was

implemented in R using "ICC.sample.size package"(Rathbone et al., 2015, Zou, 2012).

7.3.8 A record of CAP-RA Quality Control - TJC and SJC quality control

7.3.8.1 Introduction

Quality control is an essential part of any research project. Quality control defines policies and strategies that ensure that the integrity, quality, and reliability of study data are maintained throughout the life course of a research project. Quality control is a crucial facilitator of research rigour and one of the pillars of research integrity. The following describes a quality control project undertaken during the preparation of the CAP-RA study in the assessment of tender and swollen joint counts.

7.3.8.2 Method

Consecutive patients with rheumatoid arthritis in one outpatient clinic at the Sherwood Forest NHS Trust were assessed for tender and swollen joint count in a training exercise as part of the preparation for joint assessment in the CAP-RA study. All patients approached consented to participate in the training exercise. No patient identifiable patient information was collected, and verbal consent was obtained.

Tender and swollen joint counts in 28 joints, wrist, shoulder, knee, hip and the metacarpophalangeal (MCP) and proximal interphalangeal joints (PIP) of both limbs were assessed for tenderness and swelling by an experienced rheumatologist and clinical assessor, and myself in turn. Tenderness was elicited by gently pressing on each of the aforelisted

joints with the thumb and index finger while asking the patient if he/she felt pain. Joint swelling is described as swelling of the soft tissues around the joint detectable along the joint margins; it was observed by a visual inspection of each joint and assessment of the index joint for fluctuance. The first patient was used to demonstrate the appropriate technique. I assessed the patients first, while my supervisor confirmed my findings by independently assessing each patient in the format described above. A count of tender and swollen joints elicited by each assessor, was recorded and used for statistical analysis.

7.3.8.1.1 Statistical Analysis

A two-way random effect absolute agreement intraclass correlation (ICC) model was used to measure the inter-rater reliability (rater 1 – myself, rater 2 – my supervisor DAW). ICC values range from 0 – 1, indicating no correlation and perfect correlation, respectively. Values less than 0.5 indicated poor reliability, while values between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 indicated moderate, good, and excellent reliability, respectively(Koo and Li, 2016). Analysis was conducted in STATA® 16 SE.

7.3.8.3 Results

A total of 13 participants were assessed. Table 7.1 shows the mean tender and swollen joint counts for both raters. There seems to be a more considerable variation in mean SJC between the two raters, 3 (2.5) and 0.8 (0.9), compared to TJC 10.5 (7.5) and 11.4(8.2).

Table 7.1: Descriptive statistics of TJC and SJC for each rater

Mean (SD) of total TJC and SJC in 28 joints	N=13

Rater	TJC	SJC
Onosi (Rater 1)	10.5 (7.5)	3 (2.5)
David (Rater 2)	11.4 (8.2)	0.8 (0.9)

TJC – Tender Joint Count, SJC – Swollen Joint Count

Table 7.2 below shows the results of the two-way random effect absolute agreement ICC. Based on *a priori* agreed cut-offs, there was excellent reliability in the TJC assessment between both raters, ICC = 0.97 (CI: 0.90 – 0.99). However, there was poor agreement in the assessment of SJC, ICC=0.04 (CI: 0.00 – 0.58).

Table 7.2: Results of the Two-way random effect interclass correlation coefficient (ICC)

Joint Count	ICC	CI	
TJC (n=11)	0.97	0.90	0.99
SJC (n=12)	0.04	0.00	0.58

TJC – Tender Joint Count, SJC – Swollen Joint Count, ICC – Interclass Correlation Coefficient, CI – 95% Confidence Interval

7.3.8.4 Discussion

The above quality control project provides an objective exercise conducted in the preparation of the CAP-RA study. This exercise was conducted to assess the reliability of the researcher's joint assessment with that of an experienced clinician after the first teaching session. Findings showed excellent tender joint count reliability among researchers after one teaching session and little reliability between raters in the swollen joint count assessment.

Large variability in joint assessment has been described in the literature; however, standardization training has been shown to significantly reduce this variability (Grunke et al., 2010, Klinkhoff et al., 1988), particularly with

lower ICCs in SJC than TJCs (Pincus, 2008). Following standardisation of technique, SJC and TJC demonstrate good reliability among experienced assessors and remain valid tools to assess joint pain and swelling in rheumatoid arthritis(Scott and Houssien, 1996).

The excellent reliability in TJC compared to SJC possibly resulted from reliance on the patient rather than the assessor to elicit the presence/absence of pain, and this reliance did not change between assessors. SJC, on the other hand, relies more on the assessor's experience and, therefore, would require more demonstrations and practice for the inexperienced.

Clearly, the exercise demonstrates the need for quality control in research to ensure data quality and reliability during a study. This ensures that results obtained and inferences deduced from studies are valid. This reliability study highlighted the need for more training and practice sessions before actual patient assessment in the study.

The Covid-19 pandemic precluded more demonstration sessions, and thus, a log of progress made by the researcher could not be assessed. Further quality control projects planned to assess QST measurements and joint ultrasound assessments could not be conducted within the timeframe of the thesis and are therefore not presented.

7.3.8.5 Conclusion

Quality control is an important aspect of any research project, and objective measures of some quality control indicators are useful at the

early stages of a research study to document and monitor progress and highlight areas for improvement.

7.4 Discussion

7.4.1 Overview of study aims.

CAP-RA questionnaire is a brief questionnaire that encompasses traits associated with central sensitisation in one tool. This is a prospective study that primarily aims to validate CAP-RA as a means to identify people with abnormal central processing in RA. Additionally, the study aims to investigate the association between central sensitisation and fatigue using different QST modalities.

The study is based on the premise that persistent pain and fatigue in RA in some people are caused by abnormal central processing and that targeted adjuvant interventions provided to this specific population would significantly improve health outcomes in these individuals. If translated into clinical use, CAP-RA could help to assign people with RA to their optimal treatment in a more efficient manner. Of note is that although the tool was developed in a knee pain population to measure central pain sensitisation, data from the literature have described a direct relationship between central sensitisation and fatigue (Druce and McBeth, 2019).

In this study, the CAP-RA questionnaire will undergo rigorous psychometric testing to establish it as a validated tool to measure central pain mechanisms in individuals with RA. Pain catastrophising, depression, anxiety, neuropathic pain, pain distribution, sleep disturbance, fatigue, and depression have been found to represent a

single ‘central mechanisms of pain’ trait in a population with knee osteoarthritis(Akin-Akinyosoye et al., 2018). The CAP-RA study will establish the utility of this tool to capture abnormal central processing in another musculoskeletal disorder, and possibly initiate validity research across a range of conditions. This study is also designed to discover if other measures of abnormal central processing (other QST modalities) can predict future pain or fatigue and hence validate my underlying mechanistic hypothesis.

A validated instrument such as the BRAFs, a multidimensional fatigue measuring instrument, would permit a more granular examination of multiple dimensions of fatigue, unlike the unidimensional SF36/12 Vitality used in previous sections of this thesis. In addition, inflammation, specifically synovitis, is a known cause of RA pain; however, the relationship with fatigue is not consistent when utilizing laboratory measures of inflammation such as CRP and ESR. Therefore, in addition to ESR and CRP, the contribution of inflammation will be objectively measured using ultrasound scans to ensure documentation with sufficient detail and that its contribution to fatigue and pain prognosis is included within my analyses (Danielsen, 2018).

7.4.2 Strengths and limitations

Although a brief questionnaire may be acceptable to patients and feasible in busy clinical practice or large epidemiological studies, its interpretation is inevitably limited by its self-reported nature and even validated self-report instruments are prone to recall bias.

Recall bias is a type of information bias that occurs where there is differential recall of exposures. It is frequently illustrated in case controlled studies where a disease status influences the ability to recall previous exposures(Barry, 1996, Weinstock et al., 1991, Gibbons et al., 1993). Recall bias can also occur in studies such as CAP-RA where PROM reporting may be affected by the participants' current condition(Snyder and White, 1982, Blaney, 1986). Recall bias could lead to differential misclassification, which may produce biased measures of effect. In CAPRA, validated instruments were used, in addition to fatigue and pain weekly diaries, which had shorter recall periods to mitigate the effect of recall bias (Althubaiti, 2016).

Traits assessed with the CAP-RA questionnaire are multifaceted and subjective, precluding objective measurement. I attempt to address this by testing CAP-RA against QST as a proxy for central sensitisation. However, there is no gold standard for measuring CS, which is itself a multidimensional construct, and my previous study did not consistently demonstrate a statistically significant relationship between a QST modality (PPT) and fatigue. In this study, I utilise a brief QST procedure, which has been used in other research studies to measure central sensitisation, comprising both static (PPT) and dynamic (TS and CPM) test modalities(Georgopoulos et al., 2019). Each QST modality might have different measurement properties and indicate different aspects of CS, including spinal processing of nociceptive signals and descending modulation, but it is possible that these modalities may not capture the contribution of abnormal central processing to fatigue. Further research

using other measures of abnormal processing, such as functional MRI scans, may be indicated.

Only two time points will be measured, and this study was not designed to measure the ability of CAP-RA to predict persistent fatigue in RA beyond three months. This study will be conducted within a restricted geographical region (within the East Midlands of the UK) with people with active rheumatoid arthritis measured by DAS28. These findings might have limited generalizability outside this population because the East Midlands contains some of the most deprived areas in the UK, and socioeconomic disparities have been shown to negatively affect outcomes in inflammatory rheumatic diseases (Harrison et al., 2009, Harrison et al., 2005, Izadi et al., 2021, Yang et al., 2015). More bias may also occur with the weekly SMS text messages, as a subgroup of participants opting to join this section of the study may not be representative. However, usage of SMS texting has previously been shown to be widely used in older populations with arthritis(Blake et al., 2015). This study is reported in line with SPIRIT guidelines, as guidelines for cohort study protocols are not fully developed(Mahajan et al., 2020).

7.5 Conclusion

Pain and fatigue are persistent symptoms experienced by people with RA, even in well-controlled inflammatory diseases. Persistent pain and fatigue may be caused by mechanisms within the central nervous system linked to central sensitisation. This study seeks to validate a simple self-reported questionnaire to identify central mechanisms in the clinical

setting to facilitate the identification of individuals who may benefit from adjuvant treatments alongside those aiming to reduce joint inflammation.

7.6 Covid 19 Impact

Ethics application for this study was obtained in April 2020, during the national lockdown caused by the COVID-19 pandemic. Recruitment could not commence within the timelines of this thesis due to this lockdown and the ones that followed in 2020 and 2021. At the time of drafting this thesis, recruiting is ongoing managed by other researchers employed to execute the study. Consequently, no results would be presented here, and certain aspects of the protocol were written in the future tense to reflect this.

8. General discussion

8.1 Outline

This chapter summarises findings from the preceding four chapters in the context of the overarching aim and objectives of this thesis. Strengths and limitations (in addition to those discussed in previous chapters), and important considerations in terms of study design, methods and conclusions drawn from the previous chapters are discussed here. Finally, this chapter explores the implications of study findings and possible directions for future research.

8.2 Overview of thesis aim and objectives.

Fatigue may be the most bothersome symptom for many people with RA and other musculoskeletal diseases. Despite breakthroughs in inflammatory RA management and control, fatigue remains an unmet need from the patient's perspective. There has been an increase in RA fatigue-related research in the past decade. However, knowledge of the mechanisms of fatigue with translational value remains elusive. People with long-standing (established) RA have been studied extensively, but recently, with the advent of new treatment strategies and a proactive rather than reactive approach to RA management (Treat to Target strategy), a population of people with early, well-controlled inflammatory disease is presented. The overarching aim of this thesis, therefore, was to increase our understanding of fatigue in RA by unpicking some theorised associations between fatigue and other symptoms in early RA.

This thesis presented two observational studies (cross-sectional and longitudinal) in an early RA population. Findings from these studies led to the third observational study in the general population, and finally, the development of a study protocol for an observational cohort study of fatigue and pain in RA.

8.3 Summary of Results

More than 50% of people with RA report significant fatigue early in the disease course. RA fatigue is related to patient-reported outcome measures such as pain, mental health, and functional ability, with mental health as the most important determinant of fatigue. Markers of inflammation and disease activity were not significantly associated with fatigue. Interestingly, the disease duration and age at onset were not significantly associated with fatigue in early RA.

Exploring the longitudinal course of fatigue across the early RA cohort revealed that fatigue levels remained stable throughout the course of the disease examined in the cohort. There was, on average, a statistically significant change in fatigue levels over time; however, this change was not clinically significant. Meaning that people who experienced fatigue at baseline were likely to continue to experience the same fatigue levels over the course of the disease. Baseline PROMs – pain, mental health, disability, and patient's assessment of disease activity (PGA) were associated with baseline fatigue; however, these factors did not affect the change of the fatigue trajectory, consistent with the persistence of fatigue over time.

Two subgroups were embedded in this population, with one group reporting persistent fatigue values above the normed population mean (fatigue group) and the other group (no-fatigue group). As seen in the cross-sectional analysis, sex, pain, mental health, and functional ability were associated with membership in the fatigue group. In addition, comorbidities and BMI were associated with membership in the fatigue group. A marker of inflammation (ESR) was not associated with the trajectory of fatigue or fatigue group membership.

An examination of fatigue in the general population under a hypothesis that similar factors associated with fatigue in the RA population are present in the general population revealed a similar set of PROMs associated with fatigue. Although in this same population, people with knee pain experienced more fatigue and worse outcomes, such as poorer mental health and pain, fatigue was also present in people without knee pain. Statistically significant associations with PPT (a measure of central sensitisation) were not consistently observed in subgroups of this population.

Finally, as part of a project to examine the psychometric properties of a self-report tool to measure central sensitisation in people with RA, I developed a protocol to investigate the association between fatigue and pain in RA. The study design and study preparation were prepared for this thesis. Aspects of study preparation, covering quality control and patient and public involvement, were presented in addition.

The literature search related to fatigue in RA was extended until 2022 using the same search strategy described previously. Newer studies

using different analytical approaches were generally consistent with findings from this thesis. For example, an association between fatigue and pain, functional ability, depression and sleep was found in several published studies. In most of these studies, no significant association with markers of inflammation or disease activity was elicited. A phenotype of persistent fatigue in people with early RA was also described. (Lee et al., 2020, Silva et al., 2020, Lopes et al., 2021, Rodriguez-Muguruza et al., 2022, Kozlowska et al., 2022, Lee et al., 2021, Lindqvist et al., 2022, Muller et al., 2022, Hegarty et al., 2021). The association with inflammation and disease activity, however, was not consistent across all studies, probably because studies utilising component disease activity scores found an association between inflammation/disease activity and fatigue (Gao et al., 2022, Khidir et al., 2022).

Overall, collating the results obtained from this thesis in the context of existing literature, pain, mental health, sleep problems and fatigue characterise the state of 'fibriomyalgianess' in RA, described by Wolfe et al. (Wolfe, 2009). These symptoms also typify a phenotype associated with central sensitisation(Akin-Akinyosoye et al., 2018), and comprise a symptom cluster described extensively in oncology research(Baden et al., 2020, Bjerkeset et al., 2020, Donovan and Jacobsen, 2007, Fleishman, 2004, Fox and Lyon, 2006). People with this symptom cluster report poorer quality of life and worse clinical outcomes. Conditions associated with this symptom cluster, for example, fibromyalgia, chronic fatigue syndrome, and irritable bowel syndrome, are also associated with central

sensitisation(Bourke et al., 2015, Nijs et al., 2012). Central sensitisation could be the missing link that explains the association between pain, fatigue and mental health.

As highlighted in the literature review (chapter 3) and in the preceding text, fatigue is still very much an unmet need in many people with RA. The measurement of fatigue still presents a challenge for researchers. Fatigue is persistent even in intensively treated RA, with similar presentations in other disease conditions. Although showing some promise, available treatments for fatigue are still a long way off from eliminating the problem in people with RA. For these reasons, fatigue presents numerous opportunities for further research.

8.4 Novel findings and implications

This thesis highlights the need for active intervention at presentation in addition to anti-inflammatory medication in the treatment of RA. The ERAN study demonstrated the persistent nature of fatigue in RA, which translates to undue suffering for people prescribed disease-modifying medication in the hope that fatigue and associated symptoms decrease with better control of inflammatory disease. Targeting sub-populations of individuals who present with fatigue at diagnosis may provide a more cost-effective approach to the management of RA. It could result in a better quality of life for the patient and cost savings from the undue escalation of DMARD prescriptions for symptom control.

With this study, I also propose a novel approach to the study of fatigue and highlight similarities between symptoms associated with

‘fibromyalgianess’ described in RA and a symptom cluster associated with fatigue in many other disease conditions (So et al., 2009, Laird et al., 2011, Fleishman, 2004). Given that inflammatory markers did not show a significant relationship with fatigue in RA, the direction of travel for novel treatment strategies for individuals with this symptom cluster may require alternative therapies in addition to anti-inflammatory pharmacotherapy.

Central sensitisation may be the thread that captures the presence of this symptom cluster in conditions associated with severe fatigue, with and without a clear inflammatory component and represents a potential area for further research. Although this study did not definitively prove an association between central sensitisation and fatigue in the general population and in people with knee pain, my research revealed exciting howbeit inconclusive evidence to suggest that this would be a productive avenue for further research. Identifying the centrally sensitised phenotype could present target populations for future interventions(Bailly et al., 2020).

Fatigue is a complex and subjective multidimensional construct that is difficult to quantify in a standardised measure. Encapsulated in this complexity are some factors associated with fatigue – pain, sleep, and mental health- which are multifaceted constructs with overlapping causes and consequences. Identifying people with this symptom cluster will aid patient stratification and facilitate a more personalised (patient-centred) approach to managing fatigue.

8.5 Strengths and limitations

In the context of the overall aims and objectives of this thesis, and sections of the thesis completed given the covid-19 pandemic, it is essential to note some strengths and limitations. The major strength of this study is the application of a robust and sophisticated analytical framework on large observational datasets, allowing for a contribution to the body of knowledge concerning fatigue in two musculoskeletal conditions and the general population.

This study is not without limitations. Firstly, a systematic review and meta analysis was not conducted as part of the literature review. Systematic reviews are a cornerstone of rigorous scientific evidence, allowing for the systematic, objective, reproducible and transparent examination, and synthesis of literature. In this thesis, the broad nature of the research question and consequent heterogeneity of included studies, as well as time requirements did not allow for a comprehensive systematic review process. Nevertheless, the literature search was conducted using a systematic and transparent approach.

Another key limitation, and perhaps most importantly is that this thesis comprised of observational studies, from which causal associations should be drawn with caution. Observational studies carry the risk of confounding by unmeasured variables. As highlighted in the preceding chapters, sleep, a well-documented factor associated with fatigue, was not collected in the ERAN dataset; therefore, the contribution of sleep data to fatigue trajectories could not be ascertained. Additionally,

socioeconomic status, another factor associated with fatigue in general population studies, was not available in the KPIC data. Therefore, its effect could not be determined in this study. Nevertheless, these factors are unlikely to alter the overall direction of effect in this study, as overall findings from this thesis are generally in agreement with other studies (in other disease conditions) that have utilised varying covariate sets and analytical methodology.

Another limitation is the use of retrospectively collected data. ERAN and KPIC data were analysed retrospectively in this thesis, meaning I had no direct control over data quality. Be that as it may, ERAN and KPIC data were collected, parsed, and stored according to good data management practices and have been used in several peer-reviewed scientific publications, alluding to good quality data. In addition, this thesis benefitted from conscientious data management and the use of robust statistical techniques as well as sensitivity analyses, all geared towards ensuring that results presented were of the highest standard.

Missing data is a further limitation that applies to observational studies, particularly those with a long follow-up period, such as the ERAN cohort. Data from the first three years of data collection were used in this thesis to mitigate this problem. Statistical missing data handling techniques were applied in addition to sensitivity analyses, all to ensure the conclusions reached using these data were robust.

In the development of the observational study to examine the association between fatigue and pain in RA, considerable effort was applied to ensure that standardised measures were utilised in all aspects

of the study, including hands-on training on DAS28 data collection and the QST protocol. The presence of inter-observer bias cannot be eliminated entirely.

In addition, the lack of a standardised QST protocol and ideal location to elicit sensory hypersensitivity or the type of central hypersensitivity elicited in the study (pain, pressure, heat or cold) could affect the study findings. The study protocol was designed to elicit pain hypersensitivity; therefore, inferences regarding other sensory modalities may not be made using these study findings. This also applies to the site used to elicit QST, however, QST from two different sites was used to triangulate findings.

8.6 Directions for future research

Since the inception of this project, several published studies have attempted to expand our understanding of fatigue in rheumatic diseases. At the time of drafting this thesis, there are still large areas of uncharted territory relating to fatigue and how it affects people with rheumatic diseases.

Given the aims of this thesis, literature review and studies presented in this thesis, a few areas deserving critical scrutiny are described below.

Consensus building on ideal measurement instruments for fatigue presents a research priority. Research could be in the form of Delphi study design involving experts and patients across different disease conditions associated with fatigue (Boulkedid et al., 2011, Beiderbeck et al., 2021). These studies would present opportunities to develop

universally acceptable tools to delineate the presence or absence of fatigue, and identify changes in fatigue levels over time, including ideal measurement intervals for research and clinical applications.

Additionally, research into objective measures of fatigue including biomarkers is indicated. In current research and clinical practice, fatigue is assessed using questionnaires. In a recent study, structural and diffuse tension imaging (DTI) MRI showed some superiority over clinical measures of fatigue improvement in a group of participants with RA (Goñi et al., 2022). Although these findings have not been replicated or validated in other populations, an exciting avenue for further research into objective biomarkers of fatigue is presented.

Identifying a fatigue phenotype presents another exciting area of research. This could involve prospective cohort studies of disease-free genetically mapped/deeply phenotyped individuals or retrospective studies using available databases such as the UK biobank. Machine learning techniques could be employed to predict clusters of individuals or other factors associated with the fatigue phenotype(Bycroft et al., 2018). Artificial intelligence and machine learning algorithms could also be used to predict the development of fatigue in the presence or absence of a specific disease entity thereby aiding the early identification of people more likely to develop fatigue in the presence of disease conditions.

Given the findings from this thesis, a need for longitudinal studies modelling the natural history of fatigue in RA is still indicated at a more minutiae level. These studies should include extended follow-up periods and seek to identify fatigue heterogeneity and predictors of this

heterogeneity using larger cohorts thereby facilitating complex statistical modelling strategies.

More recent studies investigating fatigue treatment modalities in RA have provided better quality evidence because these studies focused on fatigue as a primary outcome. Interventions such as physical exercise, cognitive behavioural therapy and anti-inflammatory medications have been trialled in recent studies however, there is still room for future research: moderate effect sizes, uncertain long-term effects and a population of patients who continue to experience debilitating fatigue despite these interventions remains a challenge (Yentur et al., 2021, Shen et al., 2020, Bachmair et al., 2022). This presents a challenge for the development of symptom-specific interventions; for example, a clinical trial examining the effect of an intervention on fatigue may ignore the contribution of mental health, sleep, and pain. Findings from this thesis highlight the need for complex rather than single interventions. Complex in this sense refers to the number of interacting components, for example, pain and fatigue, the degree of tailoring required to deliver the intervention, for example, cognitive behavioural therapy in addition to anti-depressive pharmacotherapy, and the number and variability of outcome measures for example pain, mental health and fatigue(Skivington et al., 2021, Craig et al., 2008). Identifying people with this symptom cluster will aid patient stratification and facilitate the design of complex interventions geared towards many symptoms in a targeted population. A feasibility study and follow-on clinical trial involving a

complex intervention targeted at people with this symptom in multiple disease conditions provides a research opportunity.

The relationship between fatigue and RA disease flare warrants investigation. Examining RA flares poses a research conundrum because RA flare is a complex construct with multifaceted manifestations associated with both objective and subjective domains, like fatigue(Bozzalla-Cassione et al., 2022).

The relationship between fatigue and central sensitisation is also an area in which further research is required. Cross-sectional and longitudinal studies in which multiple modalities quantifying central sensitisation are indicated to determine the associations with fatigue and other PROMs associated with RA.

Finally, fatigue is a cross-cutting symptom that transcends several disease conditions. A different approach to fatigue research, whereby fatigue, like pain, sleep, and mental health, is studied outside the context of a specific disease entity, is indicated. This will facilitate knowledge sharing and accelerate the development of objective tools to measure the construct and pool resources to identify specific causal mechanisms. This approach has been successfully applied to pain research with remarkable results and research output across the spectrum of translational research.

9. Conclusion

In summary, fatigue still represents an unmet need in rheumatoid arthritis and the general population. People with specific characteristics in the general population are also more likely to report fatigue in RA. These characteristics are related to a symptom cluster and may be related to central sensitisation. If correctly identified, individuals with these characteristics may present opportunities for individualised care packages and represent a study population for further research.

Fatigue as a patient-reported outcome requires further evaluation outside the context of a specific disease condition. This will encourage active knowledge sharing and facilitate a better understanding of fatigue. More studies evaluating the association between fatigue and central sensitisation are indicated.

10. References

ABHISHEK, A., DOHERTY, M., KUO, C. F., MALLEN, C. D., ZHANG, W. & GRAINGE, M. J. 2017. Rheumatoid arthritis is getting less frequent-results of a nationwide population-based cohort study. *Rheumatology (Oxford),* 56**,** 736-744.

AGGARWAL, R., RINGOLD, S., KHANNA, D., NEOGI, T., JOHNSON, S. R., MILLER, A., BRUNNER, H. I., OGAWA, R., FELSON, D., OGDIE, A., ALETAHA, D. & FELDMAN, B. M. 2015. Distinctions between diagnostic and classification criteria? *Arthritis Care Res (Hoboken),* 67**,** 891-7.

AKIN-AKINYOSOYE, K., FROWD, N., MARSHALL, L., STOCKS, J., FERNANDES, G. S., VALDES, A., MCWILLIAMS, D. F., ZHANG, W., DOHERTY, M., FERGUSON, E. & WALSH, D. A. 2018. Traits associated with central pain augmentation in the Knee Pain In the Community (KPIC) cohort. *PAIN,* 159.

AKIN-AKINYOSOYE, K., JAMES, R. J. E., MCWILLIAMS, D. F., MILLAR, B., DAS NAIR, R., FERGUSON, E. & WALSH, D. A. 2021. The Central Aspects of Pain in the Knee (CAP-Knee) questionnaire; a mixed-methods study of a self-report instrument for assessing central mechanisms in people with knee pain. *Osteoarthritis and Cartilage,* 29**,** 802-814.

AKIN-AKINYOSOYE, K., SARMANOVA, A., FERNANDES, G. S., FROWD, N., SWAITHES, L., STOCKS, J., VALDES, A., MCWILLIAMS, D. F., ZHANG, W., DOHERTY, M., FERGUSON, E. & WALSH, D. A. 2020. Baseline self-report 'central mechanisms' trait predicts persistent knee pain in the Knee Pain in the Community (KPIC) cohort. *Osteoarthritis Cartilage,* 28**,** 173-181.

ALAMANOS, Y., VOULGARI, P. V. & DROSOS, A. A. 2006. Incidence and Prevalence of Rheumatoid Arthritis, Based on the 1987 American College of Rheumatology Criteria: A Systematic Review. *Seminars in Arthritis and Rheumatism,* 36**,** 182-188.

ALETAHA, D., EBERL, G., NELL, V. P., MACHOLD, K. P. & SMOLEN, J. S. 2002. Practical progress in realisation of early diagnosis and treatment of patients with suspected rheumatoid arthritis: results from two matched questionnaires within three years. *Ann Rheum Dis,* 61**,** 630-4.

ALETAHA, D., NEOGI, T., SILMAN, A. J., FUNOVITS, J., FELSON, D. T., BINGHAM, C. O., 3RD, BIRNBAUM, N. S., BURMESTER, G. R., BYKERK, V. P., COHEN, M. D., COMBE, B., COSTENBADER, K. H., DOUGADOS, M., EMERY, P., FERRACCIOLI, G., HAZES, J. M., HOBBS, K., HUIZINGA, T. W., KAVANAUGH, A., KAY, J., KVIEN, T. K., LAING, T., MEASE, P., MÉNARD, H. A., MORELAND, L. W., NADEN, R. L., PINCUS, T., SMOLEN, J. S., STANISLAWSKA-BIERNAT, E., SYMMONS, D., TAK, P. P., UPCHURCH, K. S., VENCOVSKY, J., WOLFE, F. & HAWKER, G. 2010. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis,* 69**,** 1580-8.

ALMEIDA, C., CHOY, E. H., HEWLETT, S., KIRWAN, J. R., CRAMP, F., CHALDER, T., POLLOCK, J. & CHRISTENSEN, R. 2016. Biologic interventions for fatigue in rheumatoid arthritis. *Cochrane Database of Systematic Reviews***,** CD008334.

ALTHUBAITI, A. 2016. Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc,* 9**,** 211-7.

ALYSON, L. H., RACHEL, J., SARAH, P., JOSE, M. V. & CHRIS, S. 2012. Measures of Multimorbidity and Morbidity Burden for Use in Primary Care and Community Settings: A Systematic Review and Guide. *The Annals of Family Medicine,* 10**,** 134.

ANDERSON, J., CAPLAN, L., YAZDANY, J., ROBBINS, M. L., NEOGI, T., MICHAUD, K., SAAG, K. G., O'DELL, J. R. & KAZI, S. 2012. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. *Arthritis Care Res (Hoboken),* 64**,** 640-7.

ANDERSON, R., N; ROSENBERG, HARRY, M; 1998. Age standardization of death rates; implementation of the year 2000 standard. *In:* NATIONAL CENTER FOR HEALTH, S. (ed.) *National vital statistics report.* U.S. Dept. of the Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Health Statistics.

ARAGON, J. T., FAY, P. MICHAEL, WOLLSCHLAEGER, DANIEL; OMIDPANAH, ADAM 2020. CRAN: Epitools: Epidemiology Tools. 0.5-10.1 ed.

ARNETT, F. C., EDWORTHY, S. M., BLOCH, D. A., MCSHANE, D. J., FRIES, J. F., COOPER, N. S., HEALEY, L. A., KAPLAN, S. R., LIANG, M. H. & LUTHRA, H. S. 1988. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis & Rheumatology,* 31**,** 315-324.

ASLAM, F. & KHAN, N. A. 2018. Tools for the Assessment of Comorbidity Burden in Rheumatoid Arthritis. *Front Med (Lausanne),* 5**,** 39.

ASUERO, A. G., SAYAGO, A. & GONZÁLEZ, A. G. 2006. The Correlation Coefficient: An Overview. *Critical Reviews in Analytical Chemistry,* 36**,** 41-59.

ATKINSON, T. M., ROSENFELD, B. D., SIT, L., MENDOZA, T. R., FRUSCIONE, M., LAVENE, D., SHAW, M., LI, Y., HAY, J., CLEELAND, C. S., SCHER, H. I., BREITBART, W. S. & BASCH, E. 2011. Using confirmatory factor analysis to evaluate construct validity of the Brief Pain Inventory (BPI). *J Pain Symptom Manage,* 41**,** 558-65.

AZEN, R. & TRAXEL, N. 2009. Using Dominance Analysis to Determine Predictor Importance in Logistic Regression. *Journal of Educational and Behavioral Statistics,* 34**,** 319-347.

BACALAO, E. J., GREENE, G. J., BEAUMONT, J. L., EISENSTEIN, A., MUFTIC, A., MANDELIN, A. M., CELLA, D. & RUDERMAN, E. M. 2017. Standardizing and personalizing the treat to target (T2T) approach for rheumatoid arthritis using the Patient-Reported Outcomes Measurement Information System (PROMIS): baseline findings on patient-centered treatment priorities. *Clinical rheumatology,* 36**,** 1729-1736.

BACHMAIR, E. M., MARTIN, K., AUCOTT, L., DHAUN, N., DURES, E., EMSLEY, R., GRAY, S. R., KIDD, E., KUMAR, V., LOVELL, K., MACLENNAN, G., MCNAMEE, P., NORRIE, J., PAUL, L., PACKHAM, J., RALSTON, S. H., SIEBERT, S., WEARDEN, A., MACFARLANE, G., BASU, N. & GROUP, L. S. 2022. Remotely delivered cognitive behavioural and personalised exercise interventions for fatigue severity and impact in inflammatory rheumatic diseases (LIFT): a multicentre, randomised, controlled, open-label, parallel-group trial. *Lancet Rheumatol,* 4**,** e534-e545.

BACKHAUS, M., BURMESTER, G. R., GERBER, T., GRASSI, W., MACHOLD, K. P., SWEN, W. A., WAKEFIELD, R. J. & MANGER, B. 2001. Guidelines for musculoskeletal ultrasound in rheumatology. *Ann Rheum Dis,* 60**,** 641-9.

BACKONJA, M. M., ATTAL, N., BARON, R., BOUHASSIRA, D., DRANGHOLT, M., DYCK, P. J., EDWARDS, R. R., FREEMAN, R., GRACELY, R., HAANPAA, M. H., HANSSON, P., HATEM, S. M., KRUMOVA, E. K., JENSEN, T. S., MAIER, C., MICK, G., RICE, A. S., ROLKE, R., TREEDE, R.-D., SERRA, J., TOELLE, T., TUGNOLI, V., WALK, D., WALALCE, M. S., WARE, M., YARNITSKY, D. & ZIEGLER, D. 2013. Value of quantitative sensory testing in neurological and pain disorders: NeuPSIG consensus. *PAIN®,* 154**,** 1807-1819.

BADEN, M., LU, L., DRUMMOND, F. J., GAVIN, A. & SHARP, L. 2020. Pain, fatigue and depression symptom cluster in survivors of prostate cancer. *Support Care Cancer,* 28**,** 4813-4824.

BAEK, Y., JUNG, K., KIM, H. & LEE, S. 2020. Association between fatigue, pain, digestive problems, and sleep disturbances and individuals' health-related quality of life: a nationwide survey in South Korea. *Health and Quality of Life Outcomes,* 18**,** 159.

BAGLIONI, C., NANOVSKA, S., REGEN, W., SPIEGELHALDER, K., FEIGE, B., NISSEN, C., REYNOLDS, C. F. & RIEMANN, D. 2016. Sleep and mental disorders: A meta-analysis of polysomnographic research. *Psychol Bull,* 142**,** 969-990.

BAILLY, F., CANTAGREL, A., BERTIN, P., PERROT, S., THOMAS, T., LANSAMAN, T., GRANGE, L., WENDLING, D., DOVICO, C. & TROUVIN, A. P. 2020. Part of pain labelled neuropathic in rheumatic disease might be rather nociplastic. *RMD Open,* 6.

BARON, R. M. & KENNY, D. A. 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. 51**,** 1173-1182.

BARRY, D. 1996. Differential recall bias and spurious associations in case/control studies. *Statistics in Medicine,* 15**,** 2603-2616.

BARTLETT, S. J., ORBAI, A. M., DUNCAN, T., DELEON, E., RUFFING, V., CLEGG-SMITH, K. & BINGHAM, C. O. 2015. Reliability and validity of selected PROMIS measures in people with rheumatoid arthritis. *PLoS ONE,* 10 (9) (no pagination).

BASU, N., YANG, X., LUBEN, R. N., WHIBLEY, D., MACFARLANE, G. J., WAREHAM, N. J., KHAW, K. T. & MYINT, P. K. 2016. Fatigue is associated with excess mortality in the general population: results from the EPIC-Norfolk study. *BMC Med,* 14**,** 122.

BEIDERBECK, D., FREVEL, N., VON DER GRACHT, H. A., SCHMIDT, S. L. & SCHWEITZER, V. M. 2021. Preparing, conducting, and analyzing Delphi surveys: Cross-disciplinary practices, new directions, and advancements. *MethodsX,* 8**,** 101401.

BELZA, B. L. 1995. Comparison of self-reported fatigue in rheumatoid arthritis and controls. *J Rheumatol,* 22**,** 639-43.

BELZA, B. L., HENKE, C. J., YELIN, E. H., EPSTEIN, W. V. & GILLISS, C. L. 1993. Correlates of fatigue in older adults with rheumatoid arthritis. *Nursing research,* 42**,** 93-99.

BERGMAN, M., SHAHOURI, S., SHAHOURI, S., SHAVER, T., ANDERSON, J., WEIDENSAUL, D., BUSCH, R., WANG, S. & WOLFE, F. 2009. Is Fatigue an Inflammatory Variable in Rheumatoid Arthritis (RA)? Analyses of Fatigue in RA, Osteoarthritis, and Fibromyalgia. *The Journal of rheumatology,* 36**,** 2788-94.

BI, J. 2012. A review of statistical methods for determination of relative importance of correlated predictors and identification of drivers of consumer liking. *Journal of Sensory Studies,* 27**,** 87-101.

BINGHAM, C. O., 3RD, POHL, C., WOODWORTH, T. G., HEWLETT, S. E., MAY, J. E., RAHMAN, M. U., WITTER, J. P., FURST, D. E., STRAND,

C. V., BOERS, M. & ALTEN, R. E. 2009. Developing a standardized definition for disease "flare" in rheumatoid arthritis (OMERACT 9 Special Interest Group). *J Rheumatol,* 36**,** 2335-41.
BINGHAM, C. O., ALTEN, R., BARTLETT, S. J., BYKERK, V. P., BROOKS, P. M., CHOY, E., CHRISTENSEN, R., FURST, D. E., HEWLETT, S. E., LEONG, A., MAY, J. E., MONTIE, P. A. M., POHL, C., SANDERSON, T. C., STRAND, V. & WOODWORTH, T. G. 2011. Identifying Preliminary Domains to Detect and Measure Rheumatoid Arthritis Flares: Report of the OMERACT 10 RA Flare Workshop. *The Journal of Rheumatology,* 38**,** 1751.
BJERKESET, E., RÖHRL, K. & SCHOU-BREDAL, I. 2020. Symptom cluster of pain, fatigue, and psychological distress in breast cancer survivors: prevalence and characteristics. *Breast Cancer Res Treat,* 180**,** 63-71.
BJORNER, J. B., WALLENSTEIN, G. V., MARTIN, M. C., LIN, P., BLAISDELL-GROSS, B., TAK PIECH, C. & MODY, S. H. 2007. Interpreting score differences in the SF-36 Vitality scale: using clinical conditions and functional outcomes to define the minimally important difference. *Current Medical Research and Opinion,* 23**,** 731-739.
BLAKE, H., ROBERTS, A. L., BATT, M. E. & MOSES, J. P. 2015. Motive8!: feasibility of a text messaging intervention to promote physical activity in knee osteoarthritis. *International Journal of Sports and Exercise Medicine,* 1.
BLANEY, P. H. 1986. Affect and memory: a review. *Psychol Bull,* 99**,** 229-46.
BLOMHOFF, S., JACOBSEN, M. B., SPETALEN, S., DAHM, A. & MALT, U. F. 2000. Perceptual hyperreactivity to auditory stimuli in patients with irritable bowel syndrome. *Scand J Gastroenterol,* 35**,** 583-9.
BOULKEDID, R., ABDOUL, H., LOUSTAU, M., SIBONY, O. & ALBERTI, C. 2011. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One,* 6**,** e20476.
BOURKE, J. H., LANGFORD, R. M. & WHITE, P. D. 2015. The common link between functional somatic syndromes may be central sensitisation. *Journal of Psychosomatic Research,* 78**,** 228-236.
BOURKE, J. H., WODEHOUSE, T., CLARK, L. V., CONSTANTINOU, E., KIDD, B. L., LANGFORD, R., MEHTA, V. & WHITE, P. D. 2021. Central sensitisation in chronic fatigue syndrome and fibromyalgia; a case control study. *J Psychosom Res,* 150**,** 110624.
BOWER, J. E. 2014. Cancer-related fatigue—mechanisms, risk factors, and treatments. *Nature Reviews Clinical Oncology,* 11**,** 597-609.
BOZZALLA-CASSIONE, E., GRIGNASCHI, S., XOXI, B., LUVARO, T., GRECO, M. I., MAZZUCCHELLI, I., BUGATTI, S., MONTECUCCO, C. & MANZO, A. 2022. Insights Into the Concept of Rheumatoid Arthritis Flare. *Frontiers in Medicine,* 9.
BRAMER, W. M., DE JONGE, G. B., RETHLEFSEN, M. L., MAST, F. & KLEIJNEN, J. 2018. A systematic approach to searching: an efficient and complete method to develop literature searches. *J Med Libr Assoc,* 106**,** 531-541.
BROADBENT, D. E., COOPER, P. F., FITZGERALD, P. & PARKES, K. R. 1982. The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology,* 21**,** 1-16.
BULLOCK, J., RIZVI, S. A. A., SALEH, A. M., AHMED, S. S., DO, D. P., ANSARI, R. A. & AHMED, J. 2018. Rheumatoid Arthritis: A Brief Overview of the Treatment. *Med Princ Pract,* 27**,** 501-507.
BURACK, T. S., BURACK, W. R. & KNOWLDEN, N. F. 1983. Cancer II. Distortions in standardized rates. *J Occup Med,* 25**,** 737-44.

BYCROFT, C., FREEMAN, C., PETKOVA, D., BAND, G., ELLIOTT, L. T., SHARP, K., MOTYER, A., VUKCEVIC, D., DELANEAU, O., O'CONNELL, J., CORTES, A., WELSH, S., YOUNG, A., EFFINGHAM, M., MCVEAN, G., LESLIE, S., ALLEN, N., DONNELLY, P. & MARCHINI, J. 2018. The UK Biobank resource with deep phenotyping and genomic data. *Nature,* 562**,** 203-209.

BYKERK, V. P., LIE, E., BARTLETT, S. J., ALTEN, R., BOONEN, A., CHRISTENSEN, R., FURST, D. E., HEWLETT, S., LEONG, A. L., LYDDIATT, A., MARCH, L., MAY, J. E., MONTIE, P., ORBAI, A.-M., POHL, C., SCHOLTE VOSHAAR, M., WOODWORTH, T., BINGHAM, C. O. & CHOY, E. H. 2014. Establishing a Core Domain Set to Measure Rheumatoid Arthritis Flares: Report of the OMERACT 11 RA Flare Workshop. *The Journal of Rheumatology,* 41**,** 799.

CAGNIE, B., COPPIETERS, I., DENECKER, S., SIX, J., DANNEELS, L. & MEEUS, M. 2014. Central sensitization in fibromyalgia? A systematic review on structural and functional brain MRI. *Seminars in Arthritis and Rheumatism,* 44**,** 68-75.

CAMACHO, E. M., VERSTAPPEN, S. M. M., CHIPPING, J. & SYMMONS, D. P. M. 2013. Learned helplessness predicts functional disability, pain and fatigue in patients with recent-onset inflammatory polyarthritis. *Rheumatology (United Kingdom),* 52**,** 1233-1238.

CARANDANG, K., PYATAK, E. A. & VIGEN, C. L. P. 2016. Systematic Review of Educational Interventions for Rheumatoid Arthritis. *The American journal of occupational therapy : official publication of the American Occupational Therapy Association,* 70**,** 7006290020p1-7006290020p12.

CATHÉBRAS, P. J., ROBBINS, J. M., KIRMAYER, L. J. & HAYTON, B. C. 1992. Fatigue in primary care: prevalence, psychiatric comorbidity, illness behavior, and outcome. *J Gen Intern Med,* 7**,** 276-86.

CAVANAUGH, J. E. & NEATH, A. A. 2019. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *WIREs Computational Statistics,* 11**,** e1460.

CELLA, D., YOUNT, S., SORENSEN, M., CHARTASH, E., SENGUPTA, N. & GROBER, J. 2005. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. *The Journal of Rheumatology,* 32**,** 811.

CHAN, A.-W., TETZLAFF, J. M., GØTZSCHE, P. C., ALTMAN, D. G., MANN, H., BERLIN, J. A., DICKERSIN, K., HRÓBJARTSSON, A., SCHULZ, K. F., PARULEKAR, W. R., KRLEŽA-JERIĆ, K., LAUPACIS, A. & MOHER, D. 2013. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ : British Medical Journal,* 346**,** e7586.

CHAUFFIER, K., SALLIOT, C., BERENBAUM, F. & SELLAM, J. 2012. Effect of biotherapies on fatigue in rheumatoid arthritis: a systematic review of the literature and meta-analysis. *Rheumatology,* 51**,** 60-68.

CHEN, C.-I., WANG, L., WEI, W., YUCE, H. & PHILLIPS, K. 2018. Burden of rheumatoid arthritis among US Medicare population: co-morbidities, health-care resource utilization and costs. *Rheumatology Advances in Practice,* 2**,** rky005.

CHEUNG, K. L., TEN KLOOSTER, P. M., SMIT, C., DE VRIES, H. & PIETERSE, M. E. 2017. The impact of non-response bias due to sampling in public health studies: A comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health. *BMC Public Health,* 17**,** 276.

CHIMENTI, R. L., FREY-LAW, L. A. & SLUKA, K. A. 2018. A Mechanism-Based Approach to Physical Therapist Management of Pain. *Phys Ther,* 98**,** 302-314.
CHOWDHURY, M. Z. I. & TURIN, T. C. 2020. Variable selection strategies and its importance in clinical prediction modelling. *Family medicine and community health,* 8.
CHOY, E. H. 2019. Effect of biologics and targeted synthetic disease-modifying anti-rheumatic drugs on fatigue in rheumatoid arthritis. *Rheumatology (United Kingdom),* 58(Supplement 5)**,** V51-V55.
CHURRUCA, K., POMARE, C., ELLIS, L. A., LONG, J. C., HENDERSON, S. B., MURPHY, L. E. D., LEAHY, C. J. & BRAITHWAITE, J. 2021. Patient-reported outcome measures (PROMs): A review of generic and condition-specific measures and a discussion of trends and issues. *Health Expect,* 2**4,** 1015-1024.
CIEZA, A., CAUSEY, K., KAMENOV, K., HANSON, S. W., CHATTERJI, S. & VOS, T. 2020. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet,* 396**,** 2006-2017.
CLELAND, C., FERGUSON, S., ELLIS, G. & HUNTER, R. F. 2018. Validity of the International Physical Activity Questionnaire (IPAQ) for assessing moderate-to-vigorous physical activity and sedentary behaviour of older adults in the United Kingdom. *BMC Med Res Methodol,* 18**,** 176.
COJOCARU, M., COJOCARU, I. M., SILOSI, I., VRABIE, C. D. & TANASESCU, R. 2010. Extra-articular Manifestations in Rheumatoid Arthritis. *Maedica (Bucur),* 5**,** 286-91.
CONNOLLY, D., FITZPATRICK, C., O'TOOLE, L., DORAN, M. & O'SHEA, F. 2015. Impact of Fatigue in Rheumatic Diseases in the Work Environment: A Qualitative Study. *Int J Environ Res Public Health,* 12**,** 13807-22.
CONOLLY, A. & DAVIES, B. 2018. Health survey for England 2017: adult and child overweight and obesity. *NHS Digital,* 1**,** 1-30.
COOPER, N. J. 2000. Economic burden of rheumatoid arthritis: a systematic review. *Rheumatology,* 39**,** 28-33.
CORFIELD, E. C., MARTIN, N. G. & NYHOLT, D. R. 2016. Co-occurrence and symptomatology of fatigue and depression. *Comprehensive Psychiatry,* 71**,** 1-10.
CORRAINI, P., OLSEN, M., PEDERSEN, L., DEKKERS, O. M. & VANDENBROUCKE, J. P. 2017. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. *Clin Epidemiol,* 9**,** 331-338.
CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I. & PETTICREW, M. 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ,* 337**,** a1655.
CRAMP, F., HEWLETT, S., ALMEIDA, C., KIRWAN, J. R., CHOY, E. H. S., CHALDER, T., POLLOCK, J. & CHRISTENSEN, R. 2013. Non-pharmacological interventions for fatigue in rheumatoid arthritis. *Cochrane Database of Systematic Reviews,* 2013(8) (no pagination).
CRIQUI, M. H., BARRETT-CONNOR, E. & AUSTIN, M. 1978. DIFFERENCES BETWEEN RESPONDENTS AND NON-RESPONDENTS IN A POPULATION-BASED CARDIOVASCULAR DISEASE STUDY1. *American Journal of Epidemiology,* 108**,** 367-372.
CROFFORD, L. J., KALOGERAS, K. T., MASTORAKOS, G., MAGIAKOU, M. A., WELLS, J., KANIK, K. S., GOLD, P. W., CHROUSOS, G. P. & WILDER, R. L. 1997. Circadian relationships between interleukin (IL)-6

and hypothalamic-pituitary-adrenal axis hormones: failure of IL-6 to cause sustained hypercortisolism in patients with early untreated rheumatoid arthritis. *J Clin Endocrinol Metab,* 82**,** 1279-83.
CROSBY, L. J. 1991. Factors which contribute to fatigue associated with rheumatoid arthritis. *Journal of advanced nursing,* 16(8)**,** 974-981.
CROSS, M., SMITH, E., HOY, D., CARMONA, L., WOLFE, F., VOS, T., WILLIAMS, B., GABRIEL, S., LASSERE, M., JOHNS, N., BUCHBINDER, R., WOOLF, A. & MARCH, L. 2014. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis,* 73**,** 1316-22.
CROWSON, C. S., MATTESON, E. L., MYASOEDOVA, E., MICHET, C. J., ERNSTE, F. C., WARRINGTON, K. J., DAVIS, J. M., HUNDER, G. G., THERNEAU, T. M. & GABRIEL, S. E. 2011. The Lifetime Risk of Adult-Onset Rheumatoid Arthritis and Other Inflammatory Autoimmune Rheumatic Diseases. *Arthritis and rheumatism,* 63**,** 633-639.
CULLEN, W., KEARNEY, Y. & BURY, G. 2002. Prevalence of fatigue in general practice. *Irish journal of medical science,* 171**,** 10-2.
CURRAN, P. J. 2003. Have Multilevel Models Been Structural Equation Models All Along? *Multivariate behavioral research,* 38**,** 529-569.
CURRAN, P. J., OBEIDAT, K. & LOSARDO, D. 2010. Twelve Frequently Asked Questions About Growth Curve Modeling. *J Cogn Dev,* 11**,** 121-136.
D'AGOSTINO, M. A., TERSLEV, L., AEGERTER, P., BACKHAUS, M., BALINT, P., BRUYN, G. A., FILIPPUCCI, E., GRASSI, W., IAGNOCCO, A., JOUSSE-JOULIN, S., KANE, D., NAREDO, E., SCHMIDT, W., SZKUDLAREK, M., CONAGHAN, P. G. & WAKEFIELD, R. J. 2017. Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce-Part 1: definition and development of a standardised, consensus-based scoring system. *RMD Open,* 3**,** e000428.
D'SOUZA, A., MEISSNER, B. L., TANG, B., MCKENZIE, R. S. & PIECH, C. T. 2010. Effectiveness of anti-tumor necrosis factor agents in the treatment of rheumatoid arthritis: observational study. *American health & drug benefits,* 3**,** 266-73.
DANIELSEN, M. A. 2018. Ultrasonography for diagnosis, monitoring and treatment of tenosynovitis in patients with rheumatoid arthritis. *Dan Med J,* 65.
DANTZER, R., O'CONNOR, J. C., FREUND, G. G., JOHNSON, R. W. & KELLEY, K. W. 2008. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nature reviews. Neuroscience,* 9**,** 46-56.
DARTEL, S., REPPING-WUTS, J., HOOGMOED, D. V., BLEIJENBERG, G., RIEL, P. & FRANSEN, J. 2013. Association between fatigue and pain in rheumatoid arthritis: does pain precede fatigue or does fatigue precede pain? *Arthritis care & research,* 65**,** 862-869.
DAVIS, M. C., ZAUTRA, A. J., YOUNGER, J., MOTIVALA, S. J., ATTREP, J. & IRWIN, M. R. 2008. Chronic stress and regulation of cellular markers of inflammation in rheumatoid arthritis: implications for fatigue. *Brain Behav Immun,* 22**,** 24-32.
DEANE, K. D., NORRIS, J. M. & HOLERS, V. M. 2010. Preclinical rheumatoid arthritis: identification, evaluation, and future directions for investigation. *Rheumatic diseases clinics of North America,* 36**,** 213-241.

DEBELL, M. & KROSNICK, J. A. 2009. Computing weights for American national election study survey data. *nes012427. Ann Arbor, MI, Palo Alto, CA: ANES Technical Report Series*.
DEIGHTON, C., O'MAHONY, R., TOSH, J., TURNER, C. & RUDOLF, M. 2009. Management of rheumatoid arthritis: summary of NICE guidance. *Bmj,* 338**,** b702.
DIEDERICHS, C., BERGER, K. & BARTELS, D. B. 2011. The measurement of multiple chronic diseases--a systematic review on existing multimorbidity indices. *J Gerontol A Biol Sci Med Sci,* 66**,** 301-11.
DINIZ, L. R., BALSAMO, S., SOUZA, T. Y. D., MUNIZ, L. F., MARTINS, W. R. & MOTA, L. M. H. D. 2017. Measuring fatigue with multiple instruments in a Brazilian cohort of early rheumatoid arthritis patients. *Revista brasileira de reumatologia,* 57**,** 431-437.
DIXON, W. G. & SYMMONS, D. P. M. 2005. Does early rheumatoid arthritis exist? *Best Practice & Research Clinical Rheumatology,* 19**,** 37-53.
DONOVAN, K. A. & JACOBSEN, P. B. 2007. Fatigue, Depression, and Insomnia: Evidence for a Symptom Cluster in Cancer. *Seminars in Oncology Nursing,* 23**,** 127-135.
DOUGADOS, M. 2016. Comorbidities in rheumatoid arthritis. *Curr Opin Rheumatol,* 28**,** 282-8.
DRUCE, K. L., BHATTACHARYA, Y., JONES, G. T., MACFARLANE, G. J. & BASU, N. 2016. Most patients who reach disease remission following anti-TNF therapy continue to report fatigue: results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis. *Rheumatology,* 55**,** 1786-1790.
DRUCE, K. L., JONES, G. T., MACFARLANE, G. J. & BASU, N. 2015a. Patients receiving anti-TNF therapies experience clinically important improvements in RA-related fatigue: results from the British Society for Rheumatology Biologics Register for Rheumatoid Arthritis. *Rheumatology,* 54**,** 964-971.
DRUCE, K. L., JONES, G. T., MACFARLANE, G. J., VERSTAPPEN, S. M. & BASU, N. 2015b. The longitudinal course of fatigue in rheumatoid arthritis: results from the Norfolk Arthritis Register. *The Journal of rheumatology,* 42**,** 2059-2065.
DRUCE, K. L. & MCBETH, J. 2019. Central sensitization predicts greater fatigue independently of musculoskeletal pain. *Rheumatology,* 58**,** 1923-1927.
DUPOND, J.-L. 2011. Fatigue in patients with rheumatic diseases. *Joint Bone Spine,* 78**,** 156-160.
DURCAN, L., WILSON, F. & CUNNANE, G. 2014. The effect of exercise on sleep and fatigue in rheumatoid arthritis: a randomized controlled study. *The Journal of rheumatology,* 41**,** 1966-1973.
DURES, E., KITCHEN, K., ALMEIDA, C., AMBLER, N., CLISS, A., HAMMOND, A., KNOPS, B., MORRIS, M., SWINKELS, A. & HEWLETT, S. 2012. "They didn't tell us, they made us work it out ourselves": Patient perspectives of a cognitive-behavioral program for rheumatoid arthritis fatigue. *Arthritis Care and Research,* 64(4)**,** 494-501.
EINARSSON, J. T., WILLIM, M., ERNESTAM, S., SAXNE, T., GEBOREK, P. & KAPETANOVIC, M. C. 2018. Prevalence of sustained remission in rheumatoid arthritis: impact of criteria sets and disease duration, a Nationwide Study in Sweden. *Rheumatology,* 58**,** 227-236.
EMERY, P., BLANCO, R., MALDONADO COCCO, J., CHEN, Y. C., GAICH, C. L., DELOZIER, A. M., DE BONO, S., LIU, J., ROONEY, T., CHANG, C. H. & DOUGADOS, M. 2017. Patient-reported outcomes from a

phase III study of baricitinib in patients with conventional synthetic DMARD-refractory rheumatoid arthritis. *RMD Open,* 3**,** e000410.
ENDERS, C. K. & TOFIGHI, D. 2007. Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods,* 12**,** 121-138.
ENGBERG, I., SEGERSTEDT, J., WALLER, G., WENNBERG, P. & ELIASSON, M. 2017. Fatigue in the general population- associations to age, sex, socioeconomic status, physical activity, sitting time and self-rated health: the northern Sweden MONICA study 2014. *BMC public health,* 17**,** 654-654.
ENGLAND, B. R., SAYLES, H., MIKULS, T. R., JOHNSON, D. S. & MICHAUD, K. 2015. Validation of the rheumatic disease comorbidity index. *Arthritis Care Res (Hoboken),* 67**,** 865-72.
FAKHOURI, W., LOPEZ-ROMERO, P., ANTONELLI, S., LOSI, S., ROGAI, V., BUDA, S., SANGIORGI, D., PERRONE, V. & DEGLI ESPOSTI, L. 2018. Treatment patterns, health care resource utilization and costs of rheumatoid arthritis patients in Italy: findings from a retrospective administrative database analysis. *Open Access Rheumatol,* 10**,** 103-111.
FAY, M. P. & FEUER, E. J. 1997. Confidence intervals for directly standardized rates: a method based on the gamma distribution. *Stat Med,* 16**,** 791-801.
FDA. 2009. *Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.* [Online]. Rockville, MD. Available: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. [Accessed 24/04/2024 2024].
FELDTHUSEN, C., BJORK, M., FORSBLAD-D'ELIA, H. & MANNERKORPI, K. 2013. Perception, consequences, communication, and strategies for handling fatigue in persons with rheumatoid arthritis of working age - A focus group study. *Clinical Rheumatology,* 32(5)**,** 557-566.
FELDTHUSEN, C., DEAN, E., FORSBLAD-D'ELIA, H. & MANNERKORPI, K. 2016a. Effects of person-centered physical therapy on fatigue-related variables in persons with rheumatoid arthritis: A randomized controlled trial. *Archives of Physical Medicine and Rehabilitation,* 97(1)**,** 26-36.
FELDTHUSEN, C., GRIMBY-EKMAN, A., FORSBLAD-D'ELIA, H., JACOBSSON, L. & MANNERKORPI, K. 2016b. Explanatory factors and predictors of fatigue in persons with rheumatoid arthritis: A longitudinal study. *J Rehabil Med,* 48**,** 469-76.
FERNANDES, G. S., SARMANOVA, A., WARNER, S., HARVEY, H., AKIN-AKINYOSOYE, K., RICHARDSON, H., FROWD, N., MARSHALL, L., STOCKS, J., HALL, M., VALDES, A. M., WALSH, D., ZHANG, W. & DOHERTY, M. 2017. Knee pain and related health in the community study (KPIC): a cohort study protocol. *BMC Musculoskeletal Disorders,* 18**,** 404.
FEROZ, A. H., ISLAM, M. N., TEN KLOOSTER, P. M., HASAN, M., RASKER, J. J. & HAQ, S. A. 2012. The Bengali Short Form-36 was acceptable, reliable, and valid in patients with rheumatoid arthritis. *J Clin Epidemiol,* 65**,** 1227-35.
FERWERDA, M., VAN BEUGEN, S., VAN MIDDENDORP, H., SPILLEKOM-VAN KOULIL, S., DONDERS, A. R. T., VISSER, H., TAAL, E., CREEMERS, M. C. W., VAN RIEL, P. C. L. M. & EVERS, A. W. M. 2017. A tailored-guided internet-based cognitive-behavioral intervention for patients with rheumatoid arthritis as an adjunct to

standard rheumatological care: Results of a randomized controlled trial. *Pain,* 158(5)**,** 868-878.
FERWERDA, M., VAN BEUGEN, S., VAN MIDDENDORP, H., VISSER, H., VONKEMAN, H., CREEMERS, M., VAN RIEL, P., KIEVIT, W. & EVERS, A. 2018. Tailored, therapist-guided internet-based cognitive behavioral therapy compared to care as usual for patients with rheumatoid arthritis: Economic evaluation of a randomized controlled trial. *Journal of Medical Internet Research,* 20.
FEWELL, Z., DAVEY SMITH, G. & STERNE, J. A. C. 2007. The Impact of Residual and Unmeasured Confounding in Epidemiologic Studies: A Simulation Study. *American Journal of Epidemiology,* 166**,** 646-655.
FINAN, P. H., GOODIN, B. R. & SMITH, M. T. 2013. The association of sleep and pain: an update and a path forward. *J Pain,* 14**,** 1539-52.
FLEISCHMANN, R., VAN VOLLENHOVEN, R. F., VENCOVSKY, J., ALTEN, R., DAVIES, O., MOUNTIAN, I., DE LONGUEVILLE, M., CARTER, D. & CHOY, E. 2017. Long-Term Maintenance of Certolizumab Pegol Safety and Efficacy, in Combination with Methotrexate and as Monotherapy, in Rheumatoid Arthritis Patients. *Rheumatology and therapy,* 4**,** 57-69.
FLEISHMAN, S. B. 2004. Treatment of symptom clusters: pain, depression, and fatigue. *J Natl Cancer Inst Monogr***,** 119-23.
FORBES, A., WHILE, A., MATHES, L. & GRIFFITHS, P. 2006. Health problems and health-related quality of life in people with multiple sclerosis. *Clin Rehabil,* 20**,** 67-78.
FOX, S. W. & LYON, D. E. 2006. Symptom clusters and quality of life in survivors of lung cancer. *Oncol Nurs Forum,* 33**,** 931-6.
FUHRER, R. & WESSELY, S. 1995. The epidemiology of fatigue and depression: a French primary-care study. *Psychol Med,* 25**,** 895-905.
GALLAND-DECKER, C., MARQUES-VIDAL, P. & VOLLENWEIDER, P. 2019. Prevalence and factors associated with fatigue in the Lausanne middle-aged population: a population-based, cross-sectional survey. *BMJ open,* 9**,** e027070.
GANDEK, B., WARE, J. E., AARONSON, N. K., APOLONE, G., BJORNER, J. B., BRAZIER, J. E., BULLINGER, M., KAASA, S., LEPLEGE, A., PRIETO, L. & SULLIVAN, M. 1998. Cross-Validation of Item Selection and Scoring for the SF-12 Health Survey in Nine Countries: Results from the IQOLA Project. *Journal of Clinical Epidemiology,* 51**,** 1171-1178.
GAO, L., SUN, Y., PAN, L., LI, Y., YUAN, J., CUI, X. & SHI, B. 2022. Current status and influencing factors of fatigue in patients with rheumatoid arthritis: A cross-sectional study in China. *International journal of nursing practice,* 28(3)**,** e12996.
GARWOOD, W. 2004. The Early Rheumatoid Arthritis Network (ERAN). *Musculoskeletal Care,* 2**,** 240-244.
GASTON-JOHANSSON, F., FALL-DICKSON, J. M., BAKOS, A. B. & KENNEDY, M. J. 1999. Fatigue, pain, and depression in pre-autotransplant breast cancer patients. *Cancer Pract,* 7**,** 240-7.
GEISSER, M. E., GLASS, J. M., RAJCEVSKA, L. D., CLAUW, D. J., WILLIAMS, D. A., KILENY, P. R. & GRACELY, R. H. 2008. A psychophysical study of auditory and pressure sensitivity in patients with fibromyalgia and healthy controls. *J Pain,* 9**,** 417-22.
GENTY, M., COMBE, B., KOSTINE, M., ARDOUIN, E., MOREL, J. & LUKAS, C. 2017. Improvement of fatigue in patients with rheumatoid arthritis treated with biologics: relationship with sleep disorders, depression

and clinical efficacy. A prospective, multicentre study. *Clin Exp Rheumatol,* 35**,** 85-92.
GEORGOPOULOS, V., AKIN-AKINYOSOYE, K., ZHANG, W., MCWILLIAMS, D. F., HENDRICK, P. & WALSH, D. A. 2019. Quantitative sensory testing and predicting outcomes for musculoskeletal pain, disability, and negative affect: a systematic review and meta-analysis. *Pain,* 160**,** 1920-1932.
GIANNOULI, V. 2017. Fatigue, Depression, and Obesity in Patients with Rheumatoid Arthritis; More Questions than Answers: Comment on the Article by Katz et al. *Arthritis care & research,* 69**,** 454-455.
GIBBONS, L. E., PONSONBY, A. L. & DWYER, T. 1993. A comparison of prospective and retrospective responses on Sudden infant death syndrome by case and control mothers. *American Journal of Epidemiology,* 137**,** 654-659.
GIBSON, L., ALAVA, M. H. & WAILOO, A. 2015. Progression of disease in people with rheumatoid arthritis treated with non biologic therapies.
GILTHORPE, M. S., DAHLY, D. L., TU, Y. K., KUBZANSKY, L. D. & GOODMAN, E. 2014. Challenges in modelling the random structure correctly in growth mixture models and the impact this has on model mixtures. *Journal of developmental origins of health and disease,* 5**,** 197-205.
GOADSBY, P. J., HOLLAND, P. R., MARTINS-OLIVEIRA, M., HOFFMANN, J., SCHANKIN, C. & AKERMAN, S. 2017. Pathophysiology of Migraine: A Disorder of Sensory Processing. *Physiol Rev,* 97**,** 553-622.
GOÑI, M., BASU, N., MURRAY, A. D. & WAITER, G. D. 2018. Neural Indicators of Fatigue in Chronic Diseases: A Systematic Review of MRI Studies. *Diagnostics (Basel),* 8.
GOÑI, M., BASU, N., MURRAY, A. D. & WAITER, G. D. 2022. Brain predictors of fatigue in rheumatoid arthritis: A machine learning study. *PLoS One,* 17**,** e0269952.
GRASSI, W., DE ANGELIS, R., LAMANNA, G. & CERVINI, C. 1998. The clinical features of rheumatoid arthritis. *Eur J Radiol,* 27 Suppl 1**,** S18-24.
GRAY, M. & NUKI, G. Audit of delay between symptom onset and commencement of disease modifying anti-rheumatic drugs (DMARDs) in patients with newly diagnosed rheumatoid arthritis referred to a hospital rheumatology unit. Rheumatology, 2001. OXFORD UNIV PRESS GREAT CLARENDON ST, OXFORD OX2 6DP, ENGLAND, 60-60.
GREENLAND, S. 1989. Modeling and variable selection in epidemiologic analysis. *American Journal of Public Health,* 79**,** 340-349.
GREENMYER, J. R., STACY, J. M., SAHMOUN, A. E., BEAL, J. R. & DIRI, E. 2020. DAS28-CRP Cutoffs for High Disease Activity and Remission Are Lower Than DAS28-ESR in Rheumatoid Arthritis. *ACR Open Rheumatol,* 2**,** 507-511.
GRIMM, K. J., RAM, N. & ESTABROOK, R. 2016. *Growth modeling: Structural equation and multilevel modeling approaches*, Guilford Publications.
GRON, K., ORNBJERG, L. M., HETLAND, M. L., ASLAM, F., KHAN, N. A. & JACOBS, J. 2014. The association of fatigue, comorbidity burden, disease activity, disability and gross domestic product in patients with rheumatoid arthritis. Results from 34 countries participating in the Quest-RA program. *Clin Exp Rheumatol,* 32**,** 869-77.

GROVES, R. & PEYTCHEVA, E. 2008. The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis. *Public Opinion Quarterly - PUBLIC OPIN QUART,* 72**,** 167-189.
GROVES, R. M., CIALDINI, R. B. & COUPER, M. P. 1992. UNDERSTANDING THE DECISION TO PARTICIPATE IN A SURVEY. *Public Opinion Quarterly,* 56**,** 475-495.
GRUENBERG, A. M., GOLDSTEIN, R. D. & PINCUS, H. A. 2005. Classification of depression: research and diagnostic criteria: DSM-IV and ICD-10. *Biology of depression: From novel insights to therapeutic strategies***,** 1-12.
GRUNKE, M., ANTONI, C. E., KAVANAUGH, A., HILDEBRAND, V., DECHANT, C., SCHETT, G., MANGER, B. & RONNEBERGER, M. 2010. Standardization of joint examination technique leads to a significant decrease in variability among different examiners. *J Rheumatol,* 37**,** 860-4.
GUYON, A., BALBO, M., MORSELLI, L. L., TASALI, E., LEPROULT, R., L'HERMITE-BALÉRIAUX, M., VAN CAUTER, E. & SPIEGEL, K. 2014. Adverse effects of two nights of sleep restriction on the hypothalamic-pituitary-adrenal axis in healthy men. *J Clin Endocrinol Metab,* 99**,** 2861-8.
HAMMAM, N., GAMAL, R. M., RASHED, A. M., ELFETOH, N. A., MOSAD, E. & KHEDR, E. M. 2020. Fatigue in Rheumatoid Arthritis Patients: Association With Sleep Quality, Mood Status, and Disease Activity. *Reumatol Clin (Engl Ed),* 16**,** 339-344.
HAMMER, H. B., MICHELSEN, B., SEXTON, J., UHLIG, T. & PROVAN, S. A. 2021. Fatigue is cross-sectionally not associated with objective assessments of inflammation, but changes in fatigue are associated with changes of disease activity assessments during biologic treatment of patients with established rheumatoid arthritis. *Clin Rheumatol,* 40**,** 1739-1749.
HANSEN, S. M., HETLAND, M. L., PEDERSEN, J., ØSTERGAARD, M., RUBAK, T. S. & BJORNER, J. B. 2016. Effect of Rheumatoid Arthritis on Longterm Sickness Absence in 1994-2011: A Danish Cohort Study. *J Rheumatol,* 43**,** 707-15.
HARRIS, E., D. JR 2001. Clinical Features of Rheumatoid Arthritis *In:* SLEDGE, S. R. E. D. H. J. C. B. (ed.) *Kelley's Textbook of Rheumatology.* 6th ed. Philadelphia: WB Saunders Company.
HARRISON, M. J., FARRAGHER, T. M., CLARKE, A. M., MANNING, S. C., BUNN, D. K. & SYMMONS, D. P. M. 2009. Association of functional outcome with both personal- and area-level socioeconomic inequalities in patients with inflammatory polyarthritis. *Arthritis Care & Research,* 61**,** 1297-1304.
HARRISON, M. J., TRICKER, K. J., DAVIES, L., HASSELL, A., DAWES, P., SCOTT, D. L., KNIGHT, S., DAVIS, M., MULHERIN, D. & SYMMONS, D. P. 2005. The relationship between social deprivation, disease outcome measures, and response to treatment in patients with stable, long-standing rheumatoid arthritis. *J Rheumatol,* 32**,** 2330-6.
HARTE, S. E., HARRIS, R. E. & CLAUW, D. J. 2018. The neurobiology of central sensitization. *Journal of Applied Biobehavioral Research,* 23**,** e12137.
HAYNES, W. 2013. Bonferroni Correction. *In:* DUBITZKY, W., WOLKENHAUER, O., CHO, K.-H. & YOKOTA, H. (eds.) *Encyclopedia of Systems Biology.* New York, NY: Springer New York.
HEGARTY, R. S. M., CONNER, T. S., STEBBINGS, S., FLETCHER, B. D., HARRISON, A. & TREHARNE, G. J. 2021. Understanding Fatigue-

Related Disability in Rheumatoid Arthritis and Ankylosing Spondylitis: The Importance of Daily Correlates. *Arthritis Care and Research,* 73(9)**,** 1282-1289.
HEIDARI, B. 2011. Rheumatoid Arthritis: Early diagnosis and treatment outcomes. *Caspian journal of internal medicine,* 2**,** 161-170.
HEWLETT, S., ALMEIDA, C., AMBLER, N., BLAIR, P. S., CHOY, E., DURES, E., HAMMOND, A., HOLLINGWORTH, W., KADIR, B., KIRWAN, J., PLUMMER, Z., ROOKE, C., THORN, J., TURNER, N. & POLLOCK, J. 2019a. Group cognitive-behavioural programme to reduce the impact of rheumatoid arthritis fatigue: The RAFT RCT with economic and qualitative evaluations. *Health Technology Assessment,* 23(57)**,** vii-129.
HEWLETT, S., ALMEIDA, C., AMBLER, N., BLAIR, P. S., CHOY, E. H., DURES, E., HAMMOND, A., HOLLINGWORTH, W., KADIR, B., KIRWAN, J. R., PLUMMER, Z., ROOKE, C., THORN, J., TURNER, N. & POLLOCK, J. 2019b. Reducing arthritis fatigue impact: Two-year randomised controlled trial of cognitive behavioural approaches by rheumatology teams (RAFT). *Annals of the Rheumatic Diseases,* 78(4)**,** 465-472.
HEWLETT, S., CARR, M., RYAN, S., KIRWAN, J., RICHARDS, P., CARR, A. & HUGHES, R. 2005. Outcomes generated by patients with rheumatoid arthritis: how important are they? *Musculoskeletal care,* 3**,** 131-42.
HEWLETT, S., CHALDER, T., CHOY, E., CRAMP, F., DAVIS, B., DURES, E., NICHOLLS, C. & KIRWAN, J. 2010. Fatigue in rheumatoid arthritis: time for a conceptual model. *Rheumatology,* 50**,** 1004-1006.
HEWLETT, S., DURES, E. & ALMEIDA, C. 2011. Measures of fatigue: Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional Questionnaire (BRAF MDQ), Bristol Rheumatoid Arthritis Fatigue Numerical Rating Scales (BRAF NRS) for severity, effect, and coping, Chalder Fatigue Questionnaire (CFQ), Checklist Individual Strength (CIS20R and CIS8R), Fatigue Severity Scale (FSS), Functional Assessment Chronic Illness Therapy (Fatigue) (FACIT-F), Multi-Dimensional Assessment of Fatigue (MAF), Multi-Dimensional Fatigue Inventory (MFI), Pediatric Quality Of Life (PedsQL) Multi-Dimensional Fatigue Scale, Profile of Fatigue (ProF), Short Form 36 Vitality Subscale (SF-36 VT), and Visual Analog Scales (VAS). *Arthritis Care Res (Hoboken),* 63 Suppl 11**,** S263-86.
HEWLETT, S., HEHIR, M. & KIRWAN, J. R. 2007. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. *Arthritis Care & Research,* 57**,** 429-439.
HEWLETT, S., KIRWAN, J., BODE, C., CRAMP, F., CARMONA, L., DURES, E., ENGLBRECHT, M., FRANSEN, J., GREENWOOD, R., HAGEL, S., VAN DE LAAR, M., MOLTO, A., NICKLIN, J., PETERSSON, I. F., REDONDO, M., SCHETT, G. & GOSSEC, L. 2018. The revised Bristol Rheumatoid Arthritis Fatigue measures and the Rheumatoid Arthritis Impact of Disease scale: validation in six countries. *Rheumatology (Oxford, England),* 57**,** 300-308.
HICKIE, I., KOSCHERA, A., HADZI-PAVLOVIC, D., BENNETT, B. & LLOYD, A. 1999. The temporal stability and co-morbidity of prolonged fatigue: a longitudinal study in primary care. *Psychological Medicine,* 29**,** 855-861.
HIFINGER, M., PUTRIK, P., RAMIRO, S., KESZEI, A. P., HMAMOUCHI, I., DOUGADOS, M., GOSSEC, L. & BOONEN, A. 2016. In rheumatoid arthritis, country of residence has an important influence on fatigue:

results from the multinational COMORA study. *Rheumatology (Oxford),* 55**,** 735-44.
HILL, J. C., DUNN, K. M., LEWIS, M., MULLIS, R., MAIN, C. J., FOSTER, N. E. & HAY, E. M. 2008. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum,* 59**,** 632-41.
HOSMER, D. W., JR., LEMESHOW, S., STURDIVANT, R. X. & HOSMER, D. W., JR. 2013. *Applied Logistic Regression,* New York, UNITED STATES, John Wiley & Sons, Incorporated.
HUMPHREYS, J. H., VERSTAPPEN, S. M. M., HYRICH, K. L., CHIPPING, J. R., MARSHALL, T. & SYMMONS, D. P. M. 2012. The incidence of rheumatoid arthritis in the UK: comparisons using the 2010 ACR/EULAR classification criteria and the 1987 ACR classification criteria. Results from the Norfolk Arthritis Register. *Annals of the Rheumatic Diseases*.
IFESEMEN, O. S., MCWILLIAMS, D. F., NORTON, S., KIELY, P. D. W., YOUNG, A. & WALSH, D. A. 2021. Fatigue in early rheumatoid arthritis: data from the Early Rheumatoid Arthritis Network. *Rheumatology***,** keab947.
INTERNATIONAL ASSOCIATION FOR THE STUDY OF PAIN (IASP). *IASP Terminology* [Online]. International Association for the Study of Pain [online]. Available: https://www.iasp-pain.org/resources/terminology/ [Accessed 27/05/2022 2022].
IRWIN, M. R., OLMSTEAD, R., CARRILLO, C., SADEGHI, N., FITZGERALD, J. D., RANGANATH, V. K. & NICASSIO, P. M. 2012. Sleep loss exacerbates fatigue, depression, and pain in rheumatoid arthritis. *Sleep,* 35**,** 537-543.
IZADI, Z., LI, J., EVANS, M., HAMMAM, N., KATZ, P., OGDIE, A., SUTER, L. G., YAZDANY, J. & SCHMAJUK, G. 2021. Socioeconomic Disparities in Functional Status in a National Sample of Patients With Rheumatoid Arthritis. *JAMA Network Open,* 4**,** e2119400-e2119400.
JAMAL, S., PATRA, K. & KEYSTONE, E. C. 2009. Adalimumab response in patients with early versus established rheumatoid arthritis: DE019 randomized controlled trial subanalysis. *Clinical Rheumatology,* 28**,** 413-419.
JENKINSON, C., LAYTE, R., JENKINSON, D., LAWRENCE, K., PETERSEN, S., PAICE, C. & STRADLING, J. 1997. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Public Health Med,* 19**,** 179-86.
JM, L. 1994. Sample Size and Item Calibration Stability. 7. Available: www.rasch.org/rmt/rmt74m.htm [Accessed 18/07/2022].
JOHNSON, J. W. & LEBRETON, J. M. 2004. History and Use of Relative Importance Indices in Organizational Research. *Organizational Research Methods,* 7**,** 238-257.
JOHNSTON, M. C., CRILLY, M., BLACK, C., PRESCOTT, G. J. & MERCER, S. W. 2019. Defining and measuring multimorbidity: a systematic review of systematic reviews. *European Journal of Public Health,* 29**,** 182-189.
JONES, B. L. & NAGIN, D. S. 2013. A note on a Stata plugin for estimating group-based trajectory models. *Sociological Methods & Research,* 42**,** 608-613.
JONES, D. E. J., GRAY, J. C. & NEWTON, J. 2009. Perceived fatigue is comparable between different disease groups. *QJM: An International Journal of Medicine,* 102**,** 617-624.

KANG, H. 2013. The prevention and handling of the missing data. *Korean journal of anesthesiology,* 64**,** 402-406.
KAPRAL, T., DERNOSCHNIG, F., MACHOLD, K. P., STAMM, T., SCHOELS, M., SMOLEN, J. S. & ALETAHA, D. 2007. Remission by composite scores in rheumatoid arthritis: are ankles and feet important? *Arthritis Research & Therapy,* 9**,** R72-R72.
KARSHIKOFF, B., SUNDELIN, T. & LASSELIN, J. 2017. Role of Inflammation in Human Fatigue: Relevance of Multidimensional Assessments and Potential Neuronal Mechanisms. *Frontiers in Immunology,* 8**,** 21.
KATZ, M. H. 2011. *Multivariable Analysis : A Practical Guide for Clinicians and Public Health Researchers,* Cambridge, UNITED KINGDOM, Cambridge University Press.
KATZ, P., MARGARETTEN, M., GREGORICH, S. & TRUPIN, L. 2018. Physical Activity to Reduce Fatigue in Rheumatoid Arthritis: A Randomized Controlled Trial. *Arthritis care & research,* 70**,** 1-10.
KATZ, P., MARGARETTEN, M., TRUPIN, L., SCHMAJUK, G., YAZDANY, J. & YELIN, E. 2016. Role of sleep disturbance, depression, obesity, and physical inactivity in fatigue in rheumatoid arthritis. *Arthritis care & research,* 68**,** 81-90.
KAY, J. & UPCHURCH, K. S. 2012. ACR/EULAR 2010 rheumatoid arthritis classification criteria. *Rheumatology,* 51**,** vi5-vi9.
KEYSTONE, E. C., TAYLOR, P. C., TANAKA, Y., GAICH, C., DELOZIER, A. M., DUDEK, A., ZAMORA, J. V., COBOS, J. A. C., ROONEY, T., BONO, S., ARORA, V., LINETZKY, B. & WEINBLATT, M. E. 2017. Patient-reported outcomes from a phase 3 study of baricitinib versus placebo or adalimumab in rheumatoid arthritis: secondary analyses from the RA-BEAM study. *Annals of the Rheumatic Diseases,* 76**,** 1853-1861.
KHIDIR, S. J. H., WOUTERS, F., VAN DER HELM-VAN MIL, A. H. M. & VAN MULLIGEN, E. 2022. The course of fatigue during the development of rheumatoid arthritis and its relation with inflammation: a longitudinal study. *Joint Bone Spine,* 89(6) (no pagination).
KILIC, L., ERDEN, A., GOSSEC, L. & KALYONCU, U. 2015. Fatigue, coping, sleeping disorders and productivity were still not frequently reported in rheumatoid arthritis trials published in 2014-2015: A systematic literature review. *Arthritis and Rheumatology. Conference: American College of Rheumatology/Association of Rheumatology Health Professionals Annual Scientific Meeting, ACR/ARHP,* 67.
KJEVERUD, A., ØSTLIE, K., SCHANKE, A.-K., GAY, C., THORESEN, M. & LERDAL, A. 2020. Trajectories of fatigue among stroke patients from the acute phase to 18 months post-injury: A latent class analysis. *PLOS ONE,* 15**,** e0231709.
KLIJN, S. L., WEIJENBERG, M. P., LEMMENS, P., VAN DEN BRANDT, P. A. & PASSOS, V. L. 2017. Introducing the fit-criteria assessment plot - A visualisation tool to assist class enumeration in group-based trajectory modelling. *Statistical methods in medical research,* 26**,** 2424-2436.
KLINKHOFF, A. V., BELLAMY, N., BOMBARDIER, C., CARETTE, S., CHALMERS, A., ESDAILE, J. M., GOLDSMITH, C., TUGWELL, P., SMYTHE, H. A. & BUCHANAN, W. W. 1988. An experiment in reducing interobserver variability of the examination for joint tenderness. *J Rheumatol,* 15**,** 492-4.
KLUSMANN, B., FLEER, J., TOVOTE, K. A., WEERSMA, R. K., VAN DULLEMEN, H. M., DIJKSTRA, G. & SCHROEVERS, M. J. 2021. Trajectories of Fatigue in Inflammatory Bowel Disease. *Inflamm Bowel Dis,* 27**,** 1919-1930.

KOENIGS, M. & GRAFMAN, J. 2009. The functional neuroanatomy of depression: distinct roles for ventromedial and dorsolateral prefrontal cortex. *Behavioural brain research,* 201**,** 239-243.
KOO, T. K. & LI, M. Y. 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med,* 15**,** 155-63.
KORPELA, M., LAASONEN, L., HANNONEN, P., KAUTIAINEN, H., LEIRISALO-REPO, M., HAKALA, M., PAIMELA, L., BLÅFIELD, H., PUOLAKKA, K. & MÖTTÖNEN, T. 2004. Retardation of joint damage in patients with early rheumatoid arthritis by initial aggressive treatment with disease-modifying antirheumatic drugs: Five-year experience from the FIN-RACo study. *Arthritis & Rheumatism,* 50**,** 2072-2081.
KORTE, S. M. & STRAUB, R. H. 2019. Fatigue in inflammatory rheumatic disorders: pathophysiological mechanisms. *Rheumatology (Oxford, England),* 58**,** v35-v50.
KOZLOWSKA, K. A., FORMANOWICZ, D. & BACZYK, G. 2022. The Link between the Demographic and Clinical Factors and Fatigue Symptoms among Rheumatoid Arthritis Patients. *International Journal of Environmental Research and Public Health,* 19(22) (no pagination).
KRÜGER, S., KHAYAT, D., HOFFMEISTER, M. & HILBERG, T. 2016. Pain thresholds following maximal endurance exercise. *European Journal of Applied Physiology,* 116**,** 535-540.
LAGE-HANSEN, P. R., CHRYSIDIS, S., LAGE-HANSEN, M., HOUGAARD, A., EJSTRUP, L. & AMRIS, K. 2016. Concomitant fibromyalgia in rheumatoid arthritis is associated with the more frequent use of biological therapy: a cross-sectional study. *Scand J Rheumatol,* 45**,** 45-48.
LAHIRI, M., MORGAN, C., SYMMONS, D. P. M. & BRUCE, I. N. 2012. Modifiable risk factors for RA: prevention, better than cure? *Rheumatology,* 51**,** 499-512.
LAIRD, B. J., SCOTT, A. C., COLVIN, L. A., MCKEON, A. L., MURRAY, G. D., FEARON, K. C. & FALLON, M. T. 2011. Pain, depression, and fatigue as a symptom cluster in advanced cancer. *J Pain Symptom Manage,* 42**,** 1-11.
LEE, A. N., BECK, C. E. & HALL, M. 2008. Rheumatoid factor and anti-CCP autoantibodies in rheumatoid arthritis: a review. *Clin Lab Sci,* 21**,** 15-8.
LEE, H. J., POK, L. S. L., NG, C. M., YAHYA, F., SOCKALINGAM, S., TEE, Y. C. & RAJA, J. 2020. Fatigue and associated factors in a multi-ethnic cohort of rheumatoid arthritis patients. *International Journal of Rheumatic Diseases,* 23(8)**,** 1088-1093.
LEE, S. Y., IBRAHIM, F., TOM, B. D. M., NIKIPHOROU, E., WILLIAMS, F. M. K., LEMPP, H. & SCOTT, D. L. 2021. Baseline predictors of remission, pain and fatigue in rheumatoid arthritis: the TITRATE trial. *Arthritis Res Ther,* 23**,** 278.
LEE, Y. C., NAPADOW, V. & LOGGIA, M. L. 2018. Editorial: Functional Connectivity: Dissecting the Relationship Between the Brain and "Pain Centralization" in Rheumatoid Arthritis. *Arthritis & rheumatology (Hoboken, N.J.),* 70**,** 977-980.
LEEB, B. F., ANDEL, I., SAUTNER, J., NOTHNAGL, T. & RINTELEN, B. 2004. The DAS28 in rheumatoid arthritis and fibromyalgia patients. *Rheumatology (Oxford),* 43**,** 1504-7.
LEEB, B. F., HAINDL, P. M., MAKTARI, A., NOTHNAGL, T. & RINTELEN, B. 2007. Disease activity score-28 values differ considerably depending on patient's pain perception and sex. *J Rheumatol,* 34**,** 2382-7.

LENNON, H., KELLY, S., SPERRIN, M., BUCHAN, I., CROSS, A. J., LEITZMANN, M., COOK, M. B. & RENEHAN, A. G. 2018. Framework to construct and interpret latent class trajectory modelling. *BMJ Open,* 8.

LERDAL, A., WAHL, A., RUSTØEN, T., HANESTAD, B. R. & MOUM, T. 2005. Fatigue in the general population: a translation and test of the psychometric properties of the Norwegian version of the fatigue severity scale. *Scand J Public Health,* 33**,** 123-30.

LEWIS, I., HACKETT, K. L., NG, W. F., ELLIS, J. & NEWTON, J. L. 2019. A two-phase cohort study of the sleep phenotype within primary Sjögren's syndrome and its clinical correlates. *Clin Exp Rheumatol,* 37 Suppl 118**,** 78-82.

LINDE, L., SØRENSEN, J., OSTERGAARD, M., HØRSLEV-PETERSEN, K. & HETLAND, M. L. 2008. Health-related quality of life: validity, reliability, and responsiveness of SF-36, 15D, EQ-5D [corrected] RAQoL, and HAQ in patients with rheumatoid arthritis. *J Rheumatol,* 35**,** 1528-37.

LINDQVIST, J., ALFREDSSON, L., KLARESKOG, L., LAMPA, J. & WESTERLIND, H. 2022. Unmet Needs in Rheumatoid Arthritis: A Subgroup of Patients With High Levels of Pain, Fatigue, and Psychosocial Distress 3 Years After Diagnosis. *ACR Open Rheumatology,* 4(6)**,** 492-502.

LITTLE, R. J. A. & RUBIN, D. B. 2014. *Statistical analysis with missing data*, John Wiley & Sons.

LLUCH, E., TORRES, R., NIJS, J. & VAN OOSTERWIJCK, J. 2014. Evidence for central sensitization in patients with osteoarthritis pain: A systematic literature review. *European Journal of Pain,* 18**,** 1367-1375.

LOGE, J. H., EKEBERG, Ø. & KAASA, S. 1998a. Fatigue in the general norwegian population: Normative data and associations. *Journal of Psychosomatic Research,* 45**,** 53-65.

LOGE, J. H., KAASA, S., HJERMSTAD, M. J. & KVIEN, T. K. 1998b. Translation and performance of the Norwegian SF-36 Health Survey in patients with rheumatoid arthritis. I. Data quality, scaling assumptions, reliability, and construct validity. *J Clin Epidemiol,* 51**,** 1069-76.

LOPES, F. H. A., FREITAS, M. V. C., DE BRUIN, V. M. S. & DE BRUIN, P. F. C. 2021. Depressive symptoms are associated with impaired sleep, fatigue, and disease activity in women with rheumatoid arthritis. *Advances in Rheumatology,* 61(1) (no pagination).

LOPES, P. C., BLOCK, P. & KÖNIG, B. 2016. Infection-induced behavioural changes reduce connectivity and the potential for disease spread in wild mice contact networks. *Sci Rep,* 6**,** 31790.

LØPPENTHIN, K., ESBENSEN, B. A., ØSTERGAARD, M., JENNUM, P., TOLVER, A., AADAHL, M., THOMSEN, T. & MIDTGAARD, J. 2015. Physical activity and the association with fatigue and sleep in Danish patients with rheumatoid arthritis. *Rheumatology International,* 35**,** 1655-1664.

LOUATI, K. & BERENBAUM, F. 2015. Fatigue in chronic inflammation - a link to pain pathways. *Arthritis Research & Therapy,* 17**,** 254.

LUÍS, M., GARCIA, S., GUIMARÃES, F., ANTÓNIO, M., FERNANDES, A. L., ARAÚJO, F., CUNHS, R., COUTO, M., PINTO, A. S., SILVA, L., CRUZ, M., SANTOS, M. J., SILVA, J. & DUARTE, C. 2020. Early retirement attributed to rheumatoid arthritis and its predictors. *Acta Reumatol Port,* 45**,** 177-182.

MAHAJAN, R., BURZA, S., BOUTER, L. M., SIJTSMA, K., KNOTTNERUS, A., KLEIJNEN, J., VAN DAEL, P. & ZEEGERS, M. P. 2020. Standardized protocol items recommendations for observational

studies (SPIROS) for observational study protocol reporting guidelines: protocol for a Delphi study. *JMIR Research Protocols,* 9**,** e17864.
MANCUSO, C. A., PAGET, S. A. & CHARLSON, M. E. 2000. Adaptations made by rheumatoid arthritis patients to continue working: a pilot study of workplace challenges and successful adaptations. *Arthritis Care Res,* 13**,** 89-99.
MANLY, C. A. & WELLS, R. S. 2015. Reporting the Use of Multiple Imputation for Missing Data in Higher Education Research. *Research in Higher Education,* 56**,** 397-409.
MAQUET, D., CROISIER, J.-L., DEMOULIN, C. & CRIELAARD, J.-M. 2004. Pressure pain thresholds of tender point sites in patients with fibromyalgia and in healthy controls. *European Journal of Pain,* 8**,** 111-117.
MARGARETTEN, M., JULIAN, L., KATZ, P. & YELIN, E. 2011. Depression in patients with rheumatoid arthritis: description, causes and mechanisms. *International journal of clinical rheumatology,* 6**,** 617-623.
MASKA, L., ANDERSON, J. & MICHAUD, K. 2011. Measures of functional status and quality of life in rheumatoid arthritis: Health Assessment Questionnaire Disability Index (HAQ), Modified Health Assessment Questionnaire (MHAQ), Multidimensional Health Assessment Questionnaire (MDHAQ), Health Assessment Questionnaire II (HAQ-II), Improved Health Assessment Questionnaire (Improved HAQ), and Rheumatoid Arthritis Quality of Life (RAQoL). *Arthritis Care & Research,* 63**,** S4-S13.
MATHIAS, K., AMARNANI, A., PAL, N., KARRI, J., ARKFELD, D., HAGEDORN, J. M. & ABD-ELSAYED, A. 2021. Chronic Pain in Patients with Rheumatoid Arthritis. *Current pain and headache reports,* 25**,** 59-59.
MAYOUX-BENHAMOU, M. Fatigue and rheumatoid arthritis. Annales de readaptation et de medecine physique, 2006. Elsevier, 385-388.
MCBETH, J., TOMENSON, B., CHEW-GRAHAM, C., MACFARLANE, G., JACKSON, J., LITTLEWOOD, A. & CREED, F. 2015. Common and unique associated factors for medically unexplained chronic widespread pain and chronic fatigue. *Journal of psychosomatic research,* 79**,** 484-491.
MCCORMICK, N., MARRA, C. A., SADATSAFAVI, M., KOPEC, J. A. & AVIÑA-ZUBIETA, J. A. 2019. Excess Productivity Costs of Systemic Lupus Erythematosus, Systemic Sclerosis, and Sjögren's Syndrome: A General Population-Based Study. *Arthritis Care Res (Hoboken),* 71**,** 142-154.
MCHORNEY, C. A., WARE, J. E., JR. & RACZEK, A. E. 1993. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care,* 31**,** 247-63.
MCWILLIAMS, D. F., DAWSON, O., YOUNG, A., KIELY, P. D. W., FERGUSON, E. & WALSH, D. A. 2019. Discrete Trajectories of Resolving and Persistent Pain in People With Rheumatoid Arthritis Despite Undergoing Treatment for Inflammation: Results From Three UK Cohorts. *J Pain,* 20**,** 716-727.
MCWILLIAMS, D. F. & WALSH, D. A. 2016. Factors predicting pain and early discontinuation of tumour necrosis factor-α-inhibitors in people with rheumatoid arthritis: results from the British society for rheumatology biologics register. *BMC Musculoskeletal Disorders,* 17**,** 337.
MCWILLIAMS, D. F., ZHANG, W., MANSELL, J. S., KIELY, P. D. W., YOUNG, A. & WALSH, D. A. 2012. Predictors of change in bodily pain

in early rheumatoid arthritis: An inception cohort study. *Arthritis Care & Research,* 64**,** 1505-1513.
MEIJER, J. M., MEINERS, P. M., HUDDLESTON SLATER, J. J. R., SPIJKERVET, F. K. L., KALLENBERG, C. G. M., VISSINK, A. & BOOTSMA, H. 2009. Health-related quality of life, employment and disability in patients with Sjögren's syndrome. *Rheumatology,* 48**,** 1077-1082.
MESSEMAKER, T. C., HUIZINGA, T. W. & KURREEMAN, F. 2015. Immunogenetics of rheumatoid arthritis: Understanding functional implications. *Journal of Autoimmunity,* 64**,** 74-81.
MIAN, A., IBRAHIM, F. & SCOTT, D. L. 2019. A systematic review of guidelines for managing rheumatoid arthritis. *BMC Rheumatol,* 3**,** 42.
MIDENFJORD, I., GRINSVALL, C., KOJ, P., CARNERUP, I., TÖRNBLOM, H. & SIMRÉN, M. 2021. Central sensitization and severity of gastrointestinal symptoms in irritable bowel syndrome, chronic pain syndromes, and inflammatory bowel disease. *Neurogastroenterology & Motility,* 33**,** e14156.
MINICHIELLO, E., SEMERANO, L. & BOISSIER, M. C. 2016. Time trends in the incidence, prevalence, and severity of rheumatoid arthritis: A systematic literature review. *Joint Bone Spine,* 83**,** 625-630.
MINNOCK, P., MCKEE, G., KELLY, A., CARTER, S. C., MENZIES, V., O'SULLIVAN, D., RICHARDS, P., NDOSI, M. & VAN EIJK HUSTINGS, Y. 2018. Nursing sensitive outcomes in patients with rheumatoid arthritis: A systematic literature review. *International journal of nursing studies,* 77**,** 115-129.
MOKKINK, L. B., DE VET, H. C. W., PRINSEN, C. A. C., PATRICK, D. L., ALONSO, J., BOUTER, L. M. & TERWEE, C. B. 2018. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res,* 27**,** 1171-1179.
MORRIS, J. K., TAN, J., FRYERS, P. & BESTWICK, J. 2018. Evaluation of stability of directly standardized rates for sparse data using simulation methods. *Population Health Metrics,* 16**,** 19.
MULLER, K., KUIPERS, J. G., WEIS, J., FISCHER, I., PUKROP, T., RUFFER, J. U. & KOLLER, M. 2022. Self-reported fatigue in patients with rheumatoid arthritis compared to patients with cancer: results from two large-scale studies. *Rheumatol Int,* 42**,** 329-339.
MURRAY, C. J., VOS, T., LOZANO, R., NAGHAVI, M., FLAXMAN, A. D., MICHAUD, C., EZZATI, M., SHIBUYA, K., SALOMON, J. A., ABDALLA, S., ABOYANS, V., ABRAHAM, J., ACKERMAN, I., AGGARWAL, R., AHN, S. Y., ALI, M. K., ALVARADO, M., ANDERSON, H. R., ANDERSON, L. M., ANDREWS, K. G., ATKINSON, C., BADDOUR, L. M., BAHALIM, A. N., BARKER-COLLO, S., BARRERO, L. H., BARTELS, D. H., BASÁÑEZ, M. G., BAXTER, A., BELL, M. L., BENJAMIN, E. J., BENNETT, D., BERNABÉ, E., BHALLA, K., BHANDARI, B., BIKBOV, B., BIN ABDULHAK, A., BIRBECK, G., BLACK, J. A., BLENCOWE, H., BLORE, J. D., BLYTH, F., BOLLIGER, I., BONAVENTURE, A., BOUFOUS, S., BOURNE, R., BOUSSINESQ, M., BRAITHWAITE, T., BRAYNE, C., BRIDGETT, L., BROOKER, S., BROOKS, P., BRUGHA, T. S., BRYAN-HANCOCK, C., BUCELLO, C., BUCHBINDER, R., BUCKLE, G., BUDKE, C. M., BURCH, M., BURNEY, P., BURSTEIN, R., CALABRIA, B., CAMPBELL, B., CANTER, C. E., CARABIN, H., CARAPETIS, J., CARMONA, L., CELLA, C., CHARLSON, F., CHEN, H., CHENG, A. T., CHOU, D., CHUGH, S. S., COFFENG, L. E., COLAN, S. D., COLQUHOUN, S., COLSON, K. E., CONDON, J., CONNOR, M. D.,

COOPER, L. T., CORRIERE, M., CORTINOVIS, M., DE VACCARO, K. C., COUSER, W., COWIE, B. C., CRIQUI, M. H., CROSS, M., DABHADKAR, K. C., DAHIYA, M., DAHODWALA, N., DAMSERE-DERRY, J., DANAEI, G., DAVIS, A., DE LEO, D., DEGENHARDT, L., DELLAVALLE, R., DELOSSANTOS, A., DENENBERG, J., DERRETT, S., DES JARLAIS, D. C., DHARMARATNE, S. D., et al. 2012. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet,* 380**,** 2197-223.

MUTLU, E. K. & OZDINCLER, A. R. 2015. Reliability and responsiveness of algometry for measuring pressure pain threshold in patients with knee osteoarthritis. *Journal of physical therapy science,* 27**,** 1961-1965.

NAGIN, D. 2005. *Group-Based Modeling of Development*.

NAGIN, D. S. 2014. Group-Based Trajectory Modeling: An Overview. *Annals of Nutrition and Metabolism,* 65**,** 205-210.

NAGIN, D. S. & ODGERS, C. L. 2010. Group-Based Trajectory Modeling in Clinical Research. *Annual Review of Clinical Psychology,* 6**,** 109-138.

NAREDO, E., D'AGOSTINO, M. A., WAKEFIELD, R. J., MÖLLER, I., BALINT, P. V., FILIPPUCCI, E., IAGNOCCO, A., KARIM, Z., TERSLEV, L., BONG, D. A., GARRIDO, J., MARTÍNEZ-HERNÁNDEZ, D. & BRUYN, G. A. 2013. Reliability of a consensus-based ultrasound score for tenosynovitis in rheumatoid arthritis. *Ann Rheum Dis,* 72**,** 1328-34.

NATIONAL INSTITUTE OF HEALTHCARE RESEARCH (NIHR). 2018. *What is public involvement in research?* [Online]. Available: https://www.invo.org.uk/find-out-more/what-is-public-involvement-in-research-2/ [Accessed 03/03/2021 2021].

NEATH, A. A. & CAVANAUGH, J. E. 2012. The Bayesian information criterion: background, derivation, and applications. *WIREs Computational Statistics,* 4**,** 199-203.

NEBLETT, R. 2018. The central sensitization inventory: A user's manual. *Journal of Applied Biobehavioral Research,* 23**,** e12123.

NEOGI, T., FREY-LAW, L., SCHOLZ, J., NIU, J., ARENDT-NIELSEN, L., WOOLF, C., NEVITT, M., BRADLEY, L. & FELSON, D. T. 2015. Sensitivity and sensitisation in relation to pain severity in knee osteoarthritis: trait or state? *Ann Rheum Dis,* 74**,** 682-8.

NHS DIGITAL. 2017. *Patients Registered at a GP Practice, September 2017* [Online]. Available: https://digital.nhs.uk/data-and-information/publications/statistical/patients-registered-at-a-gp-practice/september-2017 [Accessed].

NIELSEN, S. F., BOJESEN, S. E., SCHNOHR, P. & NORDESTGAARD, B. G. 2012. Elevated rheumatoid factor and long term risk of rheumatoid arthritis: a prospective cohort study. *BMJ : British Medical Journal,* 345.

NIJROLDER, I., VAN DER WINDT, D., DE VRIES, H. & VAN DER HORST, H. 2009. Diagnoses during follow-up of patients presenting with fatigue in primary care. *Cmaj,* 181**,** 683-7.

NIJS, J., MEEUS, M., VAN OOSTERWIJCK, J., ICKMANS, K., MOORKENS, G., HANS, G. & DE CLERCK, L. S. 2012. In the mind or in the brain? Scientific evidence for central sensitisation in chronic fatigue syndrome. *European Journal of Clinical Investigation,* 42**,** 203-212.

NIJS, J., TORRES-CUECO, R., VAN WILGEN, C. P., GIRBES, E. L., STRUYF, F., ROUSSEL, N., VAN OOSTERWIJCK, J., DAENEN, L., KUPPENS, K., VANWERWEEEN, L., HERMANS, L., BECKWEE, D., VOOGT, L., CLARK, J., MOLONEY, N. & MEEUS, M. 2014. Applying modern pain neuroscience in clinical practice: criteria for the classification of central sensitization pain. *Pain Physician,* 17**,** 447-57.

NIKIPHOROU, E., RADNER, H., CHATZIDIONYSIOU, K., DESTHIEUX, C., ZABALAN, C., VAN EIJK-HUSTINGS, Y., DIXON, W. G., HYRICH, K. L., ASKLING, J. & GOSSEC, L. 2016. Patient global assessment in measuring disease activity in rheumatoid arthritis: a review of the literature. *Arthritis Research & Therapy,* 18**,** 251.

NIKOLAUS, S., BODE, C., TAAL, E. & DE LAAR, M. A. 2013. Fatigue and factors related to fatigue in rheumatoid arthritis: a systematic review. *Arthritis care & research,* 65**,** 1128-1146.

NIKOLAUS, S., BODE, C., TAAL, E. & VAN DE LAAR, M. A. 2010. New insights into the experience of fatigue among patients with rheumatoid arthritis: a qualitative study. *Ann Rheum Dis,* 69**,** 895-7.

NISHIMURA, K., SUGIYAMA, D., KOGATA, Y., TSUJI, G., NAKAZAWA, T., KAWANO, S., SAIGO, K., MORINOBU, A., KOSHIBA, M., KUNTZ, K. M., KAMAE, I. & KUMAGAI, S. 2007. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. *Ann Intern Med,* 146**,** 797-808.

NORTON, S., SACKER, A., DIXEY, J., DONE, J., WILLIAMS, P. & YOUNG, A. 2013. Trajectories of functional limitation in early rheumatoid arthritis and their association with mortality. *Rheumatology (Oxford),* 52**,** 2016-24.

NOVELLA-NAVARRO, M., PLASENCIA-RODRÍGUEZ, C., NUÑO, L. & BALSA, A. 2021. Risk Factors for Developing Rheumatoid Arthritis in Patients With Undifferentiated Arthritis and Inflammatory Arthralgia. *Frontiers in Medicine,* 8.

O'BRIEN, R. M. 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & quantity,* 41**,** 673-690.

OFFICE OF NATIONAL STATISTICS. 2023. *nomis - official census and labour market statistics* [Online]. Office of National Statistics. Available: https://www.nomisweb.co.uk/reports/lmp/gor/2013265924/report.aspx#tabnrhi [Accessed 06/11/23 2023].

OLSEN, C. L., LIE, E., KVIEN, T. K. & ZANGI, H. A. 2016. Predictors of Fatigue in Rheumatoid Arthritis Patients in Remission or in a Low Disease Activity State. *Arthritis Care Res (Hoboken),* 68**,** 1043-8.

OMDAL, R., MELLGREN, S. I. & NORHEIM, K. B. 2019. Pain and fatigue in primary Sjögren's syndrome. *Rheumatology,* 60**,** 3099-3106.

OSMAN, A., BARRIOS, F. X., GUTIERREZ, P. M., KOPPER, B. A., MERRIFIELD, T. & GRITTMANN, L. 2000. The Pain Catastrophizing Scale: further psychometric evaluation with adult samples. *J Behav Med,* 23**,** 351-65.

OVERMAN, C. L., KOOL, M. B., DA SILVA, J. A. & GEENEN, R. 2016. The prevalence of severe fatigue in rheumatic diseases: an international study. *Clinical rheumatology,* 35**,** 409-415.

PALM, Ø. & PURINSZKY, E. 2005. Women with early rheumatoid arthritis are referred later than men. *Ann Rheum Dis,* 64**,** 1227-8.

PAN, Y., NORTON, S., GWINNUTT, J. M., KEARSLEY-FLEET, L., SYMMONS, D. P. M., LUNT, M., YOUNG, A., CONSORTIUM, B.-R. C. C., HYRICH, K. L. & VERSTAPPEN, S. M. M. 2019. Not all moderate disease is the same – Identification of disability trajectories among patients with rheumatoid arthritis and moderate disease activity. *PLOS ONE,* 14**,** e0215999.

PASEK, J. & PASEK, M. J. 2018. Package 'anesrake'. *Compr. R Arch. Netw*.

PAWLIKOWSKA, T., CHALDER, T., HIRSCH, S. R., WALLACE, P., WRIGHT, D. J. M. & WESSELY, S. C. 1994. Population based study of fatigue and psychological distress. *BMJ,* 308**,** 763.

PELFORT, X., TORRES-CLARAMUNT, R., SÁNCHEZ-SOLER, J. F., HINAREJOS, P., LEAL-BLANQUET, J., VALVERDE, D. & MONLLAU, J. C. 2015. Pressure algometry is a useful tool to quantify pain in the medial part of the knee: an intra- and inter-reliability study in healthy subjects. *Orthop Traumatol Surg Res,* 101**,** 559-63.

PIGEON, W. R., SATEIA, M. J. & FERGUSON, R. J. 2003. Distinguishing between excessive daytime sleepiness and fatigue: Toward improved detection and treatment. *Journal of Psychosomatic Research,* 54**,** 61-69.

PINALS, R. S., MASI, A. T. & LARSEN, R. A. 1981. Preliminary criteria for clinical remission in rheumatoid arthritis. *Arthritis & Rheumatism,* 24**,** 1308-15.

PINCUS, T. 2008. Limitations of a quantitative swollen and tender joint count to assess and monitor patients with rheumatoid arthritis. *Bull NYU Hosp Jt Dis,* 66**,** 216-23.

PLENGE, R. M. 2009. Rheumatoid arthritis genetics: 2009 update. *Current Rheumatology Reports,* 11**,** 351-356.

POLLARD, L., CHOY, E., GONZALEZ, J., KHOSHABA, B. & SCOTT, D. 2006. Fatigue in rheumatoid arthritis reflects pain, not disease activity. *Rheumatology,* 45**,** 885-889.

POWERS, J. H., 3RD, PATRICK, D. L., WALTON, M. K., MARQUIS, P., CANO, S., HOBART, J., ISAAC, M., VAMVAKAS, S., SLAGLE, A., MOLSEN, E. & BURKE, L. B. 2017. Clinician-Reported Outcome Assessments of Treatment Benefit: Report of the ISPOR Clinical Outcome Assessment Emerging Good Practices Task Force. *Value Health,* 20**,** 2-14.

PRIMDAHL, J., HEGELUND, A., LORENZEN, A. G., LOEPPENTHIN, K., DURES, E. & APPEL ESBENSEN, B. 2019. The Experience of people with rheumatoid arthritis living with fatigue: a qualitative metasynthesis. *BMJ Open,* 9**,** e024338.

PRINSEN, C. A. C., MOKKINK, L. B., BOUTER, L. M., ALONSO, J., PATRICK, D. L., DE VET, H. C. W. & TERWEE, C. B. 2018. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res,* 27**,** 1147-1157.

PROFILLIDIS, V. A. & BOTZORIS, G. N. 2019. Chapter 6 - Trend Projection and Time Series Methods. *In:* PROFILLIDIS, V. A. & BOTZORIS, G. N. (eds.) *Modeling of Transport Demand.* Elsevier.

PROUST-LIMA, C., PHILIPPS, V., DIAKITE, A., LIQUET, B., PROUST, M. C. & LIMA, P. 2021. Package 'lcmm'.

PROVAN, S. A., MICHELSEN, B., SEXTON, J., UHLIG, T. & HAMMER, H. B. 2020. Trajectories of fatigue in actively treated patients with established rheumatoid arthritis starting biologic DMARD therapy. *RMD Open,* 6**,** e001372.

R CORE TEAM 2021. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

RANZOLIN, A., BRENOL, J. C., BREDEMEIER, M., GUARIENTI, J., RIZZATTI, M., FELDMAN, D. & XAVIER, R. M. 2009. 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,* 61**,** 794-800.

RATHBONE, A., SHAW, S. & KUMBHARE, D. 2015. ICC. sample. size: calculation of sample size and power for ICC. *R Package Version,* 1.

RAZA, K., BUCKLEY, C. E., SALMON, M. & BUCKLEY, C. D. 2006. Treating very early rheumatoid arthritis. *Best practice & research. Clinical rheumatology,* 20**,** 849-863.

REMES, O., MENDES, J. F. & TEMPLETON, P. 2021. Biological, Psychological, and Social Determinants of Depression: A Review of Recent Literature. *Brain Sci,* 11.

REPPING-WUTS, H., UITTERHOEVE, R., VAN RIEL, P. & VAN ACHTERBERG, T. 2008. Fatigue as experienced by patients with rheumatoid arthritis (RA): a qualitative study. *Int J Nurs Stud,* 45**,** 995-1002.

REPPING-WUTS, H., VAN RIEL, P. & VAN ACHTERBERG, T. 2009. Fatigue in patients with rheumatoid arthritis: what is known and what is needed. *Rheumatology (Oxford),* 48**,** 207-9.

REPPING-WUTS, H., FRANSEN, J., VAN ACHTERBERG, T., BLEIJENBERG, G. & VAN RIEL, P. 2007. Persistent severe fatigue in patients with rheumatoid arthritis. *Journal of clinical nursing,* 16**,** 377-383.

REYES-GIBBY, C. C., ADAY, L. A., ANDERSON, K. O., MENDOZA, T. R. & CLEELAND, C. S. 2006. Pain, depression, and fatigue in community-dwelling adults with and without a history of cancer. *J Pain Symptom Manage,* 32**,** 118-28.

RICCI, J. A., CHEE, E., LORANDEAU, A. L. & BERGER, J. 2007. Fatigue in the US workforce: prevalence and implications for lost productive work time. *Journal of occupational and environmental medicine,* 49**,** 1-10.

RIDSDALE, L., EVANS, A., JERRETT, W., MANDALIA, S., OSLER, K. & VORA, H. 1993. Patients with fatigue in general practice: a prospective study. *Bmj,* 307**,** 103-6.

RIFBJERG-MADSEN, S., WÆHRENS, E. E., DANNESKIOLD-SAMSØE, B. & AMRIS, K. 2017. Psychometric properties of the painDETECT questionnaire in rheumatoid arthritis, psoriatic arthritis and spondyloarthritis: Rasch analysis and test-retest reliability. *Health Qual Life Outcomes,* 15**,** 110.

RIJK, A., SCHREURS, K. & BENSING, J. 2000. Patient Factors Related to the Presentation of Fatigue Complaints: Results from a Women's General Health Care Practice. *Women & health,* 30**,** 121-36.

ROBITZSCH, A., GRUND, S., HENKE, T. & ROBITZSCH, M. A. 2017. Package 'miceadds'. *R Package: Madison, WI, USA*.

ROCHE, P. A., KLESTOV, A. C. & HEIM, H. M. 2003. Description of stable pain in rheumatoid arthritis: a 6 year study. *The Journal of Rheumatology,* 30**,** 1733.

RODRIGUEZ-MUGURUZA, S., COMBE, B., GUILLEMIN, F., FAUTREL, B., OLIVE, A., MARSAL, S., VALERO, O. & LUKAS, C. 2022. Trajectories in early rheumatoid arthritis related fatigue over 10 years: results from the ESPOIR cohort. *Clinical and Experimental Rheumatology,* 40(7)**,** 1361-1367.

ROLKE, R., BARON, R., MAIER, C., TÖLLE, T. R., TREEDE, R. D., BEYER, A., BINDER, A., BIRBAUMER, N., BIRKLEIN, F., BÖTEFÜR, I. C., BRAUNE, S., FLOR, H., HUGE, V., KLUG, R., LANDWEHRMEYER, G. B., MAGERL, W., MAIHÖFNER, C., ROLKO, C., SCHAUB, C., SCHERENS, A., SPRENGER, T., VALET, M. & WASSERKA, B. 2006. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): Standardized protocol and reference values. *PAIN,* 123**,** 231-243.

RONGEN-VAN DARTEL, S., REPPING-WUTS, H., HOOGMOED, D. V., KNOOP, H., BLEIJENBERG, G., RIEL, P. & FRANSEN, J. 2014. Relationship between objectively assessed physical activity and fatigue in patients with rheumatoid arthritis: inverse correlation of activity and fatigue. *Arthritis care & research,* 66**,** 852-860.

RUBIN, D. B. 1976. Inference and Missing Data. *Biometrika,* 63**,** 581-592.
RUDAN, I., SIDHU, S., PAPANA, A., MENG, S. J., XIN-WEI, Y., WANG, W., CAMPBELL-PAGE, R. M., DEMAIO, A. R., NAIR, H., SRIDHAR, D., THEODORATOU, E., DOWMAN, B., ADELOYE, D., MAJEED, A., CAR, J., CAMPBELL, H., WANG, W. & CHAN, K. Y. 2015. Prevalence of rheumatoid arthritis in low- and middle-income countries: A systematic review and analysis. *J Glob Health,* 5**,** 010409.
RUSSELL, A. S. 2008. Quality-of-life assessment in rheumatoid arthritis. *Pharmacoeconomics,* 26**,** 831-46.
RUTA, D. A., HURST, N. P., KIND, P., HUNTER, M. & STUBBINGS, A. 1998. Measuring health status in British patients with rheumatoid arthritis: reliability, validity and responsiveness of the short form 36-item health survey (SF-36). *Br J Rheumatol,* 37**,** 425-36.
RYAN, P. 1999. SF36: Stata module to calculate summary statistics for the SF-36 Health Survey Instrument. *Statistical Software Components.* Boston College Department of Economics.
SANTOS, E. J., DUARTE, C., DA SILVA, J. A. & FERREIRA, R. J. 2019. The impact of fatigue in rheumatoid arthritis and the challenges of its assessment. *Rheumatology,* 58**,** v3-v9.
SARIYILDIZ, M. A., BATMAZ, I., BOZKURT, M., BEZ, Y., CETINCAKMAK, M. G., YAZMALAR, L., UCAR, D. & CELEPKOLU, T. 2014. Sleep quality in rheumatoid arthritis: relationship between the disease severity, depression, functional status and the quality of life. *Journal of clinical medicine research,* 6**,** 44-52.
SCERBO, T., COLASURDO, J., DUNN, S., UNGER, J., NIJS, J. & COOK, C. 2018. Measurement Properties of the Central Sensitization Inventory: A Systematic Review. *Pain Pract,* 18**,** 544-554.
SCHAFER, J. L. & GRAHAM, J. W. 2002. Missing data: Our view of the state of the art. *Psychological Methods,* 7**,** 147-177.
SCHIFF, M., TAKEUCHI, T., FLEISCHMANN, R., GAICH, C. L., DELOZIER, A. M., SCHLICHTING, D., KUO, W. L., WON, J. E., CARMACK, T., ROONEY, T., DUREZ, P., SHAIKH, S., HIDALGO, R. P., VAN VOLLENHOVEN, R. & ZERBINI, C. A. F. 2017. Patient-reported outcomes of baricitinib in patients with rheumatoid arthritis and no or limited prior disease-modifying antirheumatic drug treatment. *Arthritis Research and Therapy,* 19(1) (no pagination).
SCHOENBACH, V. J. & ROSAMOND, W. D. 2000. Understanding the fundamentals of epidemiology: an evolving text. *Chapel Hill: University of North Carolina***,** 129-151.
SCHWARZ, R., KRAUSS, O. & HINZ, A. 2003. Fatigue in the general population. *Onkologie,* 26**,** 140-4.
SCOTT, D. L. 2007. Early rheumatoid arthritis. *British Medical Bulletin,* 81-82**,** 97-114.
SCOTT, D. L. & HOUSSIEN, D. A. 1996. Joint assessment in rheumatoid arthritis. *Br J Rheumatol,* 35 Suppl 2**,** 14-8.
SHAPIRO, S. C. 2021. Biomarkers in Rheumatoid Arthritis. *Cureus* [Online], 13. Available: https://europepmc.org/articles/PMC8205440?pdf=render [Accessed 2021/05//].
SHEN, B., LI, Y., DU, X., CHEN, H., XU, Y., LI, H. & XU, G.-Y. 2020. Effects of cognitive behavioral therapy for patients with rheumatoid arthritis: A systematic review and meta-analysis. *Psychology, Health & Medicine,* 25**,** 1179-1191.
SHEN, J., BARBERA, J. & SHAPIRO, C. M. 2006. Distinguishing sleepiness and fatigue: focus on definition and measurement. *Sleep Medicine Reviews,* 10**,** 63-76.

SHEN, R., REN, X., JING, R., SHEN, X., CHEN, J., JU, S. & YANG, C. 2015. Rheumatoid factor, anti-cyclic citrullinated peptide antibody, C-reactive protein, and erythrocyte sedimentation rate for the clinical diagnosis of rheumatoid arthritis. *Laboratory medicine,* 46**,** 226-229.

SHIPPEE, N. D., DOMECQ GARCES, J. P., PRUTSKY LOPEZ, G. J., WANG, Z., ELRAIYAH, T. A., NABHAN, M., BRITO, J. P., BOEHMER, K., HASAN, R., FIRWANA, B., ERWIN, P. J., MONTORI, V. M. & MURAD, M. H. 2015. Patient and service user engagement in research: a systematic review and synthesized framework. *Health Expect,* 18**,** 1151-66.

SILVA, C. F. R., DUARTE, C., FERREIRA, R. J. O., SANTOS, E. & DA SILVA, J. A. P. 2020. Depression, disability and sleep disturbance are the main explanatory factors of fatigue in rheumatoid arthritis: a path analysis model. *Clinical and Experimental Rheumatology,* 38(2)**,** 314-321.

SKIVINGTON, K., MATTHEWS, L., SIMPSON, S. A., CRAIG, P., BAIRD, J., BLAZEBY, J. M., BOYD, K. A., CRAIG, N., FRENCH, D. P., MCINTOSH, E., PETTICREW, M., RYCROFT-MALONE, J., WHITE, M. & MOORE, L. 2021. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ,* 374**,** n2061.

SMITH, H. S., SMITH, A. R. & SEIDNER, P. 2011. Painful rheumatoid arthritis. *Pain Physician,* 14**,** E427-E458.

SMOLEN, J. S. & ALETAHA, D. 2011. Treating rheumatoid arthritis to target: Background and prospects. *International Journal of Advances in Rheumatology,* 9**,** 1-7.

SMOLEN, J. S., KREMER, J. M., GAICH, C. L., DELOZIER, A. M., SCHLICHTING, D. E., XIE, L., STOYKOV, I., ROONEY, T., BIRD, P., SANCHEZ BURSON, J. M., GENOVESE, M. C. & COMBE, B. 2017. Patient-reported outcomes from a randomised phase III study of baricitinib in patients with rheumatoid arthritis and an inadequate response to biological agents (RA-BEACON). *Annals of the Rheumatic Diseases,* 76(4)**,** 694-700.

SNEKKEVIK, H., ERIKSEN, H. R., TANGEN, T., CHALDER, T. & REME, S. E. 2014. Fatigue and depression in sick-listed chronic low back pain patients. *Pain Med,* 15**,** 1163-70.

SNYDER, M. & WHITE, P. 1982. Moods and memories: elation, depression, and the remembering of the events of one's life. *J Pers,* 50**,** 149-67.

SO, W. K., MARSH, G., LING, W. M., LEUNG, F. Y., LO, J. C., YEUNG, M. & LI, G. K. 2009. The symptom cluster of fatigue, pain, anxiety, and depression and the effect on the quality of life of women receiving treatment for breast cancer: a multicenter study. *Oncol Nurs Forum,* 36**,** E205-14.

SOKKA, T. 2003. Work disability in early rheumatoid arthritis. *Clin Exp Rheumatol,* 21**,** S71-4.

SOKKA, T. 2016. Go, go, GoTreatIT! *Clin Exp Rheumatol,* 34**,** S91-s95.

SOKKA, T. M., KAARELA, K., MOTTONEN, T. T. & HANNONEN, P. J. 1999. Conventional monotherapy compared to a "sawtooth" treatment strategy in the radiographic progression of rheumatoid arthritis over the first eight years. *Clin Exp Rheumatol,* 17**,** 527-32.

SOKOLOVE, J. & STRAND, V. 2010. Rheumatoid arthritis classification criteria - It's finally time to move on! *Bull NYU Hosp Jt Dis,* 68**,** 232-8.

SOLDATOS, C. R., DIKEOS, D. G. & PAPARRIGOPOULOS, T. J. 2000. Athens Insomnia Scale: validation of an instrument based on ICD-10 criteria. *Journal of Psychosomatic Research,* 48**,** 555-560.

SONI, A., BATRA, R. N., GWILYM, S. E., SPECTOR, T. D., HART, D. J., ARDEN, N. K., COOPER, C., TRACEY, I. & JAVAID, M. K. 2013. Neuropathic Features of Joint Pain: A Community-Based Study. *Arthritis & Rheumatism,* 65**,** 1942-1949.

SOPHIA, R.-H. & SKRONDAL, A. 2012. *Multilevel and longitudinal modeling using Stata*, Stata Press.

STANISZEWSKA, S., BRETT, J., SIMERA, I., SEERS, K., MOCKFORD, C., GOODLAD, S., ALTMAN, D. G., MOHER, D., BARBER, R., DENEGRI, S., ENTWISTLE, A., LITTLEJOHNS, P., MORRIS, C., SULEMAN, R., THOMAS, V. & TYSALL, C. 2017. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *Research Involvement and Engagement,* 3**,** 13.

STATA 2017. Stata Statistical Software: Release 15. 15 ed. College Station, TX: StataCorp LLC.

STATA 2019. Stata Statistical Software: Release 16. 16 ed. College Station, TX: StataCorp LLC.

STATA 2021. Stata Statistical Software : Release 17. 17 ed. College Station, TX: StataCorp LLC.

STEBBINGS, S., HERBISON, P. & DOYLE, T. Treharne., GJ, & Highton, J.(2010). A comparison of fatigue correlates in rheumatoid arthritis and osteoarthritis: disparity in associations with disability, anxiety and sleep disturbance. *Rheumatology (Oxford),* 49**,** 361-367.

STENBERG, D. 2007. Neuroanatomy and neurochemistry of sleep. *Cell Mol Life Sci,* 64**,** 1187-204.

STERN, A. F. 2014. The Hospital Anxiety and Depression Scale. *Occupational Medicine,* 64**,** 393-394.

STEVENSON, J., THOMPSON, M. J. J. & SONUGA-BARKE, E. 1996. Mental health of preschool children and their mothers in a mixed urban/rural population: III. Latent variable models. *The British Journal of Psychiatry,* 168**,** 26-32.

STEYERBERG, E. W., VICKERS, A. J., COOK, N. R., GERDS, T., GONEN, M., OBUCHOWSKI, N., PENCINA, M. J. & KATTAN, M. W. 2010. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology (Cambridge, Mass.),* 21**,** 128-138.

STOLTZFUS, J. C. 2011. Logistic Regression: A Brief Primer. *Academic Emergency Medicine,* 18**,** 1099-1104.

STRAND, V., GOSSEC, L., PROUDFOOT, C. W. J., CHEN, C. I., REANEY, M., GUILLONNEAU, S., KIMURA, T., VAN ADELSBERG, J., LIN, Y., MANGAN, E. K., VAN HOOGSTRATEN, H. & BURMESTER, G. R. 2018a. Patient-reported outcomes from a randomized phase III trial of sarilumab monotherapy versus adalimumab monotherapy in patients with rheumatoid arthritis. *Arthritis Res Ther,* 20**,** 129.

STRAND, V., KOSINSKI, M., CHEN, C. I., JOSEPH, G., RENDAS-BAUM, R., GRAHAM, N. M., VAN HOOGSTRATEN, H., BAYLISS, M., FAN, C., HUIZINGA, T. & GENOVESE, M. C. 2016a. Sarilumab plus methotrexate improves patient-reported outcomes in patients with active rheumatoid arthritis and inadequate responses to methotrexate: results of a phase III trial. *Arthritis Res Ther,* 18**,** 198.

STRAND, V., KREMER, J., WALLENSTEIN, G., KANIK, K. S., CONNELL, C., GRUBEN, D., ZWILLICH, S. H. & FLEISCHMANN, R. 2015. Effects of tofacitinib monotherapy on patient-reported outcomes in a randomized phase 3 study of patients with active rheumatoid arthritis and inadequate responses to DMARDs. *Arthritis Res Ther,* 17**,** 307.

STRAND, V., KREMER, J. M., GRUBEN, D., KRISHNASWAMI, S., ZWILLICH, S. H. & WALLENSTEIN, G. V. 2017a. Tofacitinib in Combination With Conventional Disease-Modifying Antirheumatic Drugs in Patients With Active Rheumatoid Arthritis: Patient-Reported Outcomes From a Phase III Randomized Controlled Trial. *Arthritis Care and Research,* 69(4)**,** 592-598.

STRAND, V., LEE, E. B., FLEISCHMANN, R., ALTEN, R. E., KONCZ, T., ZWILLICH, S. H., GRUBEN, D., WILKINSON, B., KRISHNASWAMI, S. & WALLENSTEIN, G. 2016b. Tofacitinib versus methotrexate in rheumatoid arthritis: patient-reported outcomes from the randomised phase III ORAL Start trial. *RMD open,* 2**,** e000308.

STRAND, V., MYSLER, E., MOOTS, R. J., WALLENSTEIN, G., DEMASI, R., GRUBEN, D., SOMA, K., IIKUNI, N. & FLEISCHMANN, R. 2018b. Tofacitinib with and without methotrexate versus adalimumab with methotrexate for the treatment of rheumatoid arthritis: patient-reported outcomes from a phase 3b/4 randomised trial. *Annals of the Rheumatic Diseases,* 77(Supplement 2)**,** 990-991.

STRAND, V., REANEY, M., CHEN, C. I., PROUDFOOT, C. W. J., GUILLONNEAU, S., BAUER, D., MANGAN, E., GRAHAM, N. M. H., VAN HOOGSTRATEN, H., LIN, Y., PACHECO-TENA, C. & FLEISCHMANN, R. 2017b. Sarilumab improves patient-reported outcomes in rheumatoid arthritis patients with inadequate response/intolerance to tumour necrosis factor inhibitors. *RMD Open,* 3(1) (no pagination).

STREINER, D. L., NORMAN, G. R. & CAIRNEY, J. 2015. *Health Measurement ScalesA practical guide to their development and use: A practical guide to their development and use,* Oxford, UK, Oxford University Press.

SYMMONS, D., BARRETT, E., BANKHEAD, C., SCOTT, D. & SILMAN, A. 1994. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *Rheumatology,* 33**,** 735-739.

TACK, B. B. 1990. Fatigue in rheumatoid arthritis: Conditions, strategies, and consequences. *Arthritis & Rheumatism,* 3**,** 65-70.

TÉLLEZ, N., RÍO, J., TINTORÉ, M., NOS, C., GALÁN, I. & MONTALBAN, X. 2006. Fatigue in multiple sclerosis persists over time. *Journal of Neurology,* 253**,** 1466-1470.

TENNANT, A. & CONAGHAN, P. G. 2007. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum,* 57**,** 1358-62.

TERWEE, C. B., PRINSEN, C. A. C., CHIAROTTO, A., WESTERMAN, M. J., PATRICK, D. L., ALONSO, J., BOUTER, L. M., DE VET, H. C. W. & MOKKINK, L. B. 2018. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res,* 27**,** 1159-1170.

THOMSEN, T., AADAHL, M., BEYER, N., HETLAND, M. L., LOPPENTHIN, K., MIDTGAARD, J., CHRISTENSEN, R., OSTERGAARD, M., JENNUM, P. J. & ESBENSEN, B. A. 2017. The efficacy of motivational counselling and SMS reminders on daily sitting time in patients with rheumatoid arthritis: A randomised controlled trial. *Annals of the Rheumatic Diseases,* 76(9)**,** 1603-1606.

THOMSEN, T., BEYER, N., AADAHL, M., HETLAND, M. L., LØPPENTHIN, K., MIDTGAARD, J. & ESBENSEN, B. A. 2015. Sedentary behaviour in patients with rheumatoid arthritis: A qualitative study. *Int J Qual Stud Health Well-being,* 10**,** 28578.

TIKLY, M. 2009. Modern management of rheumatoid arthritis—making a case for early aggressive medical treatment. *South African Family Practice,* 51**,** 284-290.
TOLEDANO, E., ORTIZ, A. M., IVORRA-CORTES, J., MONTES, N., BELTRAN, A., RODRIGUEZ-RODRIGUEZ, L., CARMONA, L. & GONZALEZ-ALVARO, I. 2018. Are rheumatologists adhering to the concepts window of opportunity and treat-to-target? Earlier and more intense disease-modifying anti-rheumatic drug treatment over time in patients with early arthritis in the PEARL study. *Clin Exp Rheumatol,* 36**,** 382-388.
TOURNADRE, A., PEREIRA, B., GOSSEC, L., SOUBRIER, M. & DOUGADOS, M. 2018. Impact of comorbidities on fatigue in rheumatoid arthritis patients: results from a nurse-led program for comorbidities management (COMEDRA). *Joint Bone Spine*.
TRIPEPI, G., JAGER, K. J., DEKKER, F. W. & ZOCCALI, C. 2010a. Selection Bias and Information Bias in Clinical Research. *Nephron Clinical Practice,* 115**,** c94-c99.
TRIPEPI, G., JAGER, K. J., DEKKER, F. W. & ZOCCALI, C. 2010b. Stratification for Confounding – Part 1: The Mantel-Haenszel Formula. *Nephron Clinical Practice,* 116**,** c317-c321.
TSENG, B. Y., GAJEWSKI, B. J. & KLUDING, P. M. 2010. Reliability, Responsiveness, and Validity of the Visual Analog Fatigue Scale to Measure Exertion Fatigue in People with Chronic Stroke: A Preliminary Study. *Stroke Research and Treatment,* 2010**,** 412964.
TSOUCALAS, G. & SGANTZOS, M. 2017. Primary Asthenic gout by Augustin-Jacob Landre-Beauvais in 1800: is this the first description of rheumatoid arthritis? *Mediterranean journal of rheumatology,* 28**,** 223-226.
TURESSON, C. 2016. Comorbidity in rheumatoid arthritis. *Swiss Med Wkly,* 146**,** w14290.
TURESSON, C. & JACOBSSON, L. T. 2004. Epidemiology of extra-articular manifestations in rheumatoid arthritis. *Scand J Rheumatol,* 33**,** 65-72.
TWISK, J. W. 2013. *Applied longitudinal data analysis for epidemiology: a practical guide*, Cambridge University Press.
UDDIN, Z. & MACDERMID, J. C. 2016. Quantitative Sensory Testing in Chronic Musculoskeletal Pain. *Pain Med,* 17**,** 1694-703.
VAN DE SCHOOT, R., SIJBRANDIJ, M., WINTER, S. D., DEPAOLI, S. & VERMUNT, J. K. 2017. The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. *Structural Equation Modeling: A Multidisciplinary Journal,* 24**,** 451-467.
VAN DER ELST, K., MATHIJSSEN, E. G. E., LANDGREN, E., BREMANDER, A., DE GROEF, A., LINDQVIST, E., NYLANDER, M., PETERS, A., VAN DEN HOOGEN, F., VAN EIJK-HUSTINGS, Y., VERHOEVEN, G., VRIEZEKOLK, J. E., WESTHOVENS, R. & LARSSON, I. 2020. What do patients prefer? A multinational, longitudinal, qualitative study on patient-preferred treatment outcomes in early rheumatoid arthritis. *RMD Open,* 6.
VAN DER LINDEN, M. P., KNEVEL, R., HUIZINGA, T. W. & VAN DER HELM-VAN MIL, A. H. 2011. Classification of rheumatoid arthritis: comparison of the 1987 American College of Rheumatology criteria and the 2010 American College of Rheumatology/European League Against Rheumatism criteria. *Arthritis Rheum,* 63**,** 37-42.
VAN DER WOUDE, D., HOUWING-DUISTERMAAT, J. J., TOES, R. E. M., HUIZINGA, T. W. J., THOMSON, W., WORTHINGTON, J., VAN DER HELM-VAN MIL, A. H. M. & DE VRIES, R. R. P. 2009. Quantitative

heritability of anti–citrullinated protein antibody–positive and anti–citrullinated protein antibody–negative rheumatoid arthritis. *Arthritis & Rheumatism,* 60**,** 916-923.
VAN HOOGMOED, D., FRANSEN, J., BLEIJENBERG, G. & VAN RIEL, P. 2010. Physical and psychosocial correlates of severe fatigue in rheumatoid arthritis. *Rheumatology,* 49**,** 1294-1302.
VAN LOON, A. J., TIJHUIS, M., PICAVET, H. S., SURTEES, P. G. & ORMEL, J. 2003. Survey non-response in the Netherlands: effects on prevalence estimates and associations. *Ann Epidemiol,* 13**,** 105-10.
VAN STEENBERGEN, H. W., TSONAKA, R., HUIZINGA, T. W. J., BOONEN, A. & VAN DER HELM-VAN MIL, A. H. M. 2015. Fatigue in rheumatoid arthritis; a persistent problem: a large longitudinal study. *RMD open,* 1**,** e000041.
VAN TUYL, L. H., SADLONOVA, M., HEWLETT, S., DAVIS, B., FLUREY, C., GOEL, N., GOSSEC, L., HEEGAARD BRAHE, C., HILL, C., HOOGLAND, W., KIRWAN, J., HETLAND, M. L., VAN SCHAARDENBURG, D., SMOLEN, J., SINGH, J., STAMM, T., VOSHAAR, M., WELLS, G. A. & BOERS, M. 2015. The patient perspective on absence of disease activity in rheumatoid arthritis: A survey to identify key domains of patient perceived remission. *Annals of the Rheumatic Diseases,* 2)**,** 435-436.
VIALA-DANTEN, M., MARTIN, S., GUILLEMIN, I. & HAYS, R. D. 2008. Evaluation of the reliability and validity of the Medical Outcomes Study sleep scale in patients with painful diabetic peripheral neuropathy during an international clinical trial. *Health Qual Life Outcomes,* 6**,** 113.
VICKERS, A. J. & ALTMAN, D. G. 2001. Statistics notes: Analysing controlled trials with baseline and follow up measurements. *Bmj,* 323**,** 1123-4.
VINCENT, A., BENZO, R. P., WHIPPLE, M. O., MCALLISTER, S. J., ERWIN, P. J. & SALIGAN, L. N. 2013. Beyond pain in fibromyalgia: insights into the symptom of fatigue. *Arthritis Res Ther,* 15**,** 221.
WABE, N., WOJCIECHOWSKI, J., WECHALEKAR, M. D., CLELAND, L. G., MCWILLIAMS, L., LEE, A., PROUDMAN, S. & WIESE, M. D. 2017. Disease activity trajectories in early rheumatoid arthritis following intensive DMARD therapy over 3 years: association with persistence to therapy. *Int J Rheum Dis,* 20**,** 1447-1456.
WALSH, D. A. & MCWILLIAMS, D. F. 2014. Mechanisms, impact and management of pain in rheumatoid arthritis. *Nat Rev Rheumatol,* 10**,** 581-92.
WALTER, M. J. M., KUIJPER, T. M., HAZES, J. M. W., WEEL, A. E. & LUIME, J. J. 2018. Fatigue in early, intensively treated and tight-controlled rheumatoid arthritis patients is frequent and persistent: a prospective study. *Rheumatology international,* 38**,** 1643-1650.
WANG, T., YIN, J., MILLER, A. H. & XIAO, C. 2017. A systematic review of the association between fatigue and genetic polymorphisms. *Brain Behav Immun,* 62**,** 230-244.
WARE, J. E. 1993. SF-36 health survey: manual and interpretation guide. *Health Institute*.
WARE, J. E., JR. & SHERBOURNE, C. D. 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care,* 30**,** 473-83.
WARE, J. E., KOSINSKI, M. & KELLER, S. D. 1996. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. *Medical Care,* 34**,** 220-233.
WARE JR, J. E. 1999. SF-36 health survey.

WATT, T., GROENVOLD, M., BJORNER, J. B., NOERHOLM, V., RASMUSSEN, N. A. & BECH, P. 2000. Fatigue in the Danish general population. Influence of sociodemographic factors and disease. *J Epidemiol Community Health,* 54**,** 827-33.
WEI, P., LU, Z. & SONG, J. 2015. Variable importance analysis: A comprehensive review. *Reliability Engineering & System Safety,* 142**,** 399-432.
WEINSTOCK, M. A., COLDITZ, G. A., WILLET, W. C., STAMPFER, M. J., ROSNER, B. & SPEIZER, F. E. 1991. Recall (report) bias and reliability in the retrospective assessment of melanoma risk. *American Journal of Epidemiology,* 133**,** 240-245.
WEISMAN, M. H. 2011. *Rheumatoid Arthritis*, Oxford University Press.
WEISMAN, M. H. 2014. *Rheumatoid Arthritis,* Cary, United States, Oxford University Press.
WEST, B. T., WELCH, K. B. & GALECKI, A. T. 2006. *Linear mixed models: a practical guide using statistical software*, Chapman and Hall/CRC.
WESTHOFF, G., DÖRNER, T. & ZINK, A. 2012. Fatigue and depression predict physician visits and work disability in women with primary Sjögren's syndrome: results from a cohort study. *Rheumatology (Oxford),* 51**,** 262-9.
WESTHOVENS, R., VAN DER ELST, K., MATTHYS, A., TRAN, M. & GILLOTEAU, I. 2014. Sleep problems in patients with rheumatoid arthritis. *J Rheumatol,* 41**,** 31-40.
WHIBLEY, D., ALKANDARI, N., KRISTENSEN, K., BARNISH, M., RZEWUSKA, M., DRUCE, K. L. & TANG, N. K. Y. 2019. Sleep and Pain: A Systematic Review of Studies of Mediation. *The Clinical Journal of Pain,* 35.
WHITE, I. R., ROYSTON, P. & WOOD, A. M. 2011. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med,* 30**,** 377-99.
WHITEHEAD, L. 2009. The Measurement of Fatigue in Chronic Illness: A Systematic Review of Unidimensional and Multidimensional Fatigue Measures. *Journal of Pain and Symptom Management,* 37**,** 107-128.
WIDDIFIELD, J., BERNATSKY, S., PATERSON, J. M., TOMLINSON, G., TU, K., KURIYA, B., THORNE, J. C., POPE, J. E., HOLLANDS, S. & BOMBARDIER, C. 2015. Trends in Excess Mortality Among Patients With Rheumatoid Arthritis in Ontario, Canada. *Arthritis Care Res (Hoboken),* 67**,** 1047-53.
WILLIAMS, D. A. 2018. Phenotypic Features of Central Sensitization. *Journal of applied biobehavioral research,* 23**,** e12135.
WILLIAMSON, R. J., PURCELL, S., STERNE, A., WESSELY, S., HOTOPF, M., FARMER, A. & SHAM, P. C. 2005. The relationship of fatigue to mental and physical health in a community sample. *Social Psychiatry and Psychiatric Epidemiology,* 40**,** 126-132.
WOLFE, F. 2009. Fibromyalgianess. *Arthritis Care & Research,* 61**,** 715-716.
WOLFE, F., CLAUW, D. J., FITZCHARLES, M.-A., GOLDENBERG, D. L., HÄUSER, W., KATZ, R. L., MEASE, P. J., RUSSELL, A. S., RUSSELL, I. J. & WALITT, B. 2016. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. *Seminars in Arthritis and Rheumatism,* 46**,** 319-329.
WOLFE, F., CLAUW, D. J., FITZCHARLES, M. A., GOLDENBERG, D. L., KATZ, R. S., MEASE, P., RUSSELL, A. S., RUSSELL, I. J., WINFIELD, J. B. & YUNUS, M. B. 2010. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and

measurement of symptom severity. *Arthritis Care Res (Hoboken),* 62**,** 600-10.
WORTINGER, L. A., GLENNE ØIE, M., ENDESTAD, T. & BRUUN WYLLER, V. 2017. Altered right anterior insular connectivity and loss of associated functions in adolescent chronic fatigue syndrome. *PloS one,* 12**,** e0184325-e0184325.
WU, Y., LEVIS, B., SUN, Y., HE, C., KRISHNAN, A., NEUPANE, D., BHANDARI, P. M., NEGERI, Z., BENEDETTI, A. & THOMBS, B. D. 2021. Accuracy of the Hospital Anxiety and Depression Scale Depression subscale (HADS-D) to screen for major depression: systematic review and individual participant data meta-analysis. *BMJ,* 373**,** n972.
WUKICH, D. K., SAMBENEDETTO, T. L., MOTA, N. M., SUDER, N. C. & ROSARIO, B. L. 2016. Correlation of SF-36 and SF-12 Component Scores in Patients With Diabetic Foot Disease. *The Journal of Foot and Ankle Surgery,* 55**,** 693-696.
WYSOCKI, T., OLESIŃSKA, M. & PARADOWSKA-GORYCKA, A. 2020. Current Understanding of an Emerging Role of HLA-DRB1 Gene in Rheumatoid Arthritis-From Research to Clinical Practice. *Cells,* 9**,** 1127.
YANG, G., BYKERK, V. P., BOIRE, G., HITCHON, C. A., THORNE, J. C., TIN, D., HARAOUI, B., KEYSTONE, E. C. & POPE, J. E. 2015. Does socioeconomic status affect outcomes in early inflammatory arthritis? Data from a canadian multisite suspected rheumatoid arthritis inception cohort. *J Rheumatol,* 42**,** 46-54.
YARNITSKY, D., BOUHASSIRA, D., DREWES, A. M., FILLINGIM, R. B., GRANOT, M., HANSSON, P., LANDAU, R., MARCHAND, S., MATRE, D., NILSEN, K. B., STUBHAUG, A., TREEDE, R. D. & WILDER-SMITH, O. H. 2015. Recommendations on practice of conditioned pain modulation (CPM) testing. *Eur J Pain,* 19**,** 805-6.
YENTUR, S. B., ATAS, N., OZTURK, M. A. & OSKAY, D. 2021. Comparison of the effectiveness of pilates exercises, aerobic exercises, and pilates with aerobic exercises in patients with rheumatoid arthritis. *Irish Journal of Medical Science,* 190(3)**,** 1027-1034.
YUAN, Y. & HUNT, R. H. 2009. Systematic Reviews: The Good, the Bad, and the Ugly. *Official journal of the American College of Gastroenterology | ACG,* 104**,** 1086-1092.
YUNUS, M. B. 2007. Fibromyalgia and Overlapping Disorders: The Unifying Concept of Central Sensitivity Syndromes. *Seminars in Arthritis and Rheumatism,* 36**,** 339-356.
ZATARAIN, E. & STRAND, V. 2006. Monitoring disease activity of rheumatoid arthritis in clinical practice: contributions from clinical trials. *Nat Clin Pract Rheumatol,* 2**,** 611-8.
ZEIDLER, H. 2011. How can misclassification be prevented when using the 2010 American College of Rheumatology/European League Against Rheumatism rheumatoid arthritis classification criteria? Comment on the article by van der Linden et al. *Arthritis Rheum,* 63**,** 2544-6; author reply 2456.
ZEIDLER, H. 2012. The need to better classify and diagnose early and very early rheumatoid arthritis. *J Rheumatol,* 39**,** 212-7.
ZIGMOND, A. S. & SNAITH, R. P. 1983. The hospital anxiety and depression scale. *Acta Psychiatr Scand,* 67**,** 361-70.
ZOU, G. Y. 2012. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Statistics in Medicine,* 31**,** 3972-3981.

11. Appendices

11.1 Publications from PhD

RHEUMATOLOGY

Rheumatology 2022;61:3737–3745
https://doi.org/10.1093/rheumatology/keab947
Advance Access publication 27 December 2021

Original article

Fatigue in early rheumatoid arthritis: data from the Early Rheumatoid Arthritis Network

Onosi Sylvia Ifesemen [1,2], Daniel Frederick McWilliams [1,2,3], Sam Norton [4], Patrick D. W. Kiely [5], Adam Young [6] and David Andrew Walsh[1,2,3,7]

Abstract

Objectives. Fatigue is a disabling symptom in people with RA. This study aims to describe the prevalence, risk factors and longitudinal course of fatigue in early RA.

Methods. Demographic, clinical, quality of life (QoL), comorbidities and laboratory data were from the Early RA Network (ERAN), a UK multicentre inception cohort of people with RA. Fatigue was measured using the vitality subscale of the 36-item Short Form Health Survey, where higher values represent better QoL. Baseline prevalences of fatigue classifications were age and sex standardized. Linear regression, hierarchical growth curve modelling and group-based trajectory modelling (GBTM) were utilized.

Results. At baseline ($n = 1236$, 67% female, mean age 57 years), the mean vitality was 41 (S.D. 11) and disease duration was 11 months (interquartile range 7–18). Age- and sex-standardized prevalence rates of fatigue and severe fatigue were 44% (95% CI 39, 50) and 19% (95% CI 15, 23), respectively. Fatigue changed little over 3 years and five measurement occasions $\beta = -0.13$ (95% CI -0.23, -0.02). GBTM identified two subgroups, which we named 'Fatigue' (53%) and 'No-fatigue' (47%). Female sex, worse pain, mental health and functional ability were associated with greater fatigue and predicted Fatigue group membership (area under the receiver operating characteristics curve = 0.81). Objective measures of inflammation—swollen joint count and ESR—were not significantly associated with fatigue.

Conclusions. Fatigue is prevalent and persistent in early RA. Diverse characteristics indicative of central mechanisms are associated with persistent fatigue. Management of fatigue might require interventions targeted at central mechanisms in addition to inflammatory disease modification. People who require such interventions might be identified at presentation with early RA.

Key words: fatigue, rheumatoid arthritis, trajectories, inflammation, central mechanisms

Rheumatology key messages

- Fatigue is prevalent and remains stable regardless of improvements in inflammatory disease activity.
- People with persistent fatigue could be identified early, representing targets for fatigue-lowering interventions.
- Traits indicative of central mechanisms are associated with persistent fatigue in early RA.

Ifesemen et al. BMC Rheumatology (2021) 5:23
https://doi.org/10.1186/s41927-021-00187-2

BMC Rheumatology

STUDY PROTOCOL Open Access

Central Aspects of Pain in Rheumatoid Arthritis (CAP-RA): protocol for a prospective observational study

Onosi S. Ifesemen[1,2*], Daniel F. McWilliams[1,2,3], Eamonn Ferguson[2,4], Richard Wakefield[5], Kehinde Akin-Akinyosoye[6], Deborah Wilson[7], Dorothy Platts[8], Susan Ledbury[8] and David A. Walsh[1,2,3,7]

Abstract

Background: Pain and fatigue are persistent problems in people with rheumatoid arthritis. Central sensitisation (CS) may contribute to pain and fatigue, even when treatment has controlled inflammatory disease. This study aims to validate a self-report 8-item questionnaire, the Central Aspects of Pain in Rheumatoid Arthritis (CAP-RA) questionnaire, developed to measure central pain mechanisms in RA, and to predict patient outcomes and response to treatment. A secondary objective is to explore mechanisms linking CS, pain and fatigue in people with RA.

Methods/design: This is a prospective observational cohort study recruiting 250 adults with active RA in secondary care. The CAP-RA questionnaire, demographic data, medical history, and patient reported outcome measures (PROMs) of traits associated with central sensitization will be collected using validated questionnaires. Quantitative sensory testing modalities of pressure pain detection thresholds, temporal summation and conditioned pain modulation will be indices of central sensitization, and blood markers, swollen joints and ultrasound scans will be indices of inflammation. Primary data collection will be at baseline and 12 weeks. The test-retest reliability of CAP-RA questionnaire will be determined 1 week after the baseline visit. Pain and fatigue data will be collected weekly via text messages for 12 weeks. CAP-RA psychometric properties, and predictive validity for outcomes at 3 months will be evaluated.

Discussion: This study will validate a simple self-report questionnaire against psychophysical indices of central sensitization and patient reported outcome measures of traits associated with CS in a population of individuals with active RA. The application of this instrument in the clinical environment could provide a mechanism-based stratification tool to facilitate the provision of targeted therapy to individuals with pain and fatigue in RA, alongside treatments that target joint inflammation.

Trial registration: Clinicaltrials.gov NCT04515589. Date of registration 17 August 2020.

Keywords: Rheumatoid arthritis, Pain, Central sensitization, Fatigue, Inflammation

* Correspondence: mbxosi@nottingham.ac.uk
[1]Division of Rheumatology, Orthopaedics and Dermatology, School of Medicine, University of Nottingham, Nottingham, UK
[2]Pain Centre Versus Arthritis, University of Nottingham, Nottingham, UK
Full list of author information is available at the end of the article

11.2 Literature review protocol

Literature review using a systematic approach - protocol

Aim	Describe the published literature of fatigue in RA
Objectives	Describe the evidence base available around the epidemiology of fatigue in RA
	Identify risk factors and associated with fatigue in RA
	Describe fatigue in RA from a patient's perspective
	Identify fatigue measurement tools in RA
	Describe treatment modalities for fatigue in RA
	Identify gaps in the evidence base around fatigue in RA
Criteria for including studies in the review	
Population	Adults aged 18 years and over with RA. No gender or geographical restrictions will be applied; however, only papers written in English will be reviewed.
Exposure	Rheumatoid Arthritis
Comparison or control group	N/A
Outcome of interest	Fatigue measured using a validated instrument, or a binary yes/no variable
Study designs	Published research articles using cross-sectional, cohort study designs or RCTs. Qualitative and quantitative studies will be reviewed.
Criteria for excluding studies not covered in inclusion criteria	Opinion papers or press releases and conference abstracts
Search methods	Searches will be conducted according to the methodology described by Bramer et al. Three databases - Embase, Medline and PsycInfo will be searched. Reference lists of full text will be searched
Methods of review	
Details of methods	Titles and abstracts will be exported into reference managing software and screened by a single reviewer. Full texts will be screened by a single reviewer as well.
Quality assessment	Quality assessment will be carried out using study design specific quality assessment tools
Data extraction	Excel sheet summarising salient points of article
Narrative Synthesis	Evidence will be integrated using narrative synthesis based on specific objectives.
Presentation of results	Results will be presented in the literature review section of a PhD thesis

11.3 Databases

11.3.1 ERAN Data Dictionary

variable name	*New variable name*	*Description*	*Corresponding Questionnaire (baseline or follow up)*	*Corresponding Questionnaire Question (Text)*		*value label*	*variable label*
cenid	**cenid**	Unique identifier of participants in the cohort. It is made up of the patient number (after removing one of the zeros) and the centre number.	Baseline	Derived		CENID--unique case ID	
CEN	**cen**	This is the centre number. ERAN data was collected by Rheumatologists based at these centres.	Baseline	Centre No		centre number	
PTNO			Baseline	Patient No		unique ID	
ID			Baseline	-		person number (within each centre)	
DO1V	**dov_0**	Date of the first visit in the DD MMM YR format	Baseline	Date of Visit		date of visit 1	
AGE	**age_0**	Age at first visit, Derived from Date of Birth which was not provided in master dataset	Baseline	Derived		age	
SEX	**sex**	Gender of participant	Baseline	Sex	SEX	sex	
ETHNIC	**ethnic_full**	Ethnicity of participant	**Baseline**		ETHNIC	ethnicity	
ETHOTH						ethoth	
VISIT	**visit_0**	Visit number -				visit number	
DOV						date of visit 1	

variable name	*New variable name*	*Description*	*Corresponding Questionnaire (baseline or follow up)*	*Corresponding Questionnaire Question (Text)*	*value label*	*variable label*
PTLOST	**px_outcome_0**	Outcome of participant detaills (lost to follow op , wrong diagnosis etc) Variable name - Details of participant outcome	**Not sure Baseline)**	**(** Not found	PTLOST	patient lost details
DIAG	**final_diagnosis_0**	Final_diagnosis final diagnosis & other outcomes	**Not sure Baseline)**	**(** Not found		diagnosis
WEIGHT	**weight_0**	Weight in KGs (Rounded up/down to the nearest KG)	**Baseline**	Weight in Kg		weight
HEIGHT	**height_0**	Height in CM (Rounded up/down to the nearest CM)	**Baseline**	Height in CM		height
SMOKE	**smoke_0**	Smoking history	**Baseline**	Smoking	SMOKE	smoke
SMOKEYRS	**smoke_yrs_0**	No of years smoked	**Baseline**	Not found (Probably the same as pack years)		Smoke years
PACKYRS	**packyrs_0**	Pack years	**Baseline**	Derived		Pack years
EMPLOY					EMPLOY	Employment status
RAREL	**ra_retired_0**	RA related retirement/unemployment	**Baseline**	Derived	RAREL	RA related retirement/unemployment
CURRJOB	**currjob_0**	Current job	**Baseline**	If employed current job or previously employed, please specify job		
CURRJOBL	**currjobl_0**	Some sort of job code	**Baseline**	Derived		

variable name	*New variable name*	*Description*	*Corresponding Questionnaire (baseline or follow up)*	*Corresponding Questionnaire Question (Text)*	*value label*	*variable label*
BENEFITS	**benefit_0**	Receiving benefits/allowances	**Baseline**	Receiving any benefits/ allowances?	BENEFITS	benefits-Y/N
EXRA	**exra_0**	Extra-articular Rheumatoid Arthritis	**Baseline**	Has the patient any Extra Articular Disease?	EXRA	extra-articular
SJS	**sjs_0**	Sjogren's syndrome	**Baseline**	Sjogren's	SJS	Sjogren’s syndrome
SCL	**scl_0**	Scleritis	**Baseline**	Scleritis/ Episcleritis	SCL	Scleritis/Episcleritis
LUNG	**lung_0**	Lung	**Baseline**	Lung	LUNG	lung complication
FEL	**fel_0**	Felty’s	**Baseline**	Felty’s	FEL	Felty’s
SYS	**sys_0**	Systemic Vasculitis	**Baseline**	Systemic Vasculitis	SYS	Systemic Vasculitis
NEU	**neu_0**	Neuropathy/ Myopathy	**Baseline**	Neuropathy/ Myopathy	NEU	Neuropathy/Myopathy
RAY	**ray_0**	Raynaud's	**Baseline**	Raynaud's	RAY	Raynaud’s
DIGVAS	**digvas_0**	Digital Vasculitis	**Baseline**	Digital Vasculitis	DIGVAS	Digital Vasculitis
OEXRA	**oexra_0**	Other extra articular disease	**Baseline**	Other	OEXRA	Other ExRA
DOEXRA	**doexra_0**	Other extra articular disease	**Baseline**	Other?		Details other ExRA
DAA	**daa_0**	Assessor's global assessment of Disease Activity	**Baseline**	How active do you feel your Patient's Rheumatoid Arthritis is today?	DAA	Assessor's global assessment of Disease Activity
DPGA	**dpga_0**	Date of patient’s global assessment of disease activity	**Baseline**	Patient's date of PGA Global Assessment of Disease Activity. Date Performed:		
PGA		Patient's assessment of disease activity in mm ranging from 0 (not active at all) to 100 (extremely active)	**Baseline**	Considering all the ways your arthritis affects you, how active do you feel your Rheumatoid Arthritis has been in the past week?		patient's global assessment of dx activity (mm)

variable name	*New variable name*	*Description*	*Corresponding Questionnaire (baseline or follow up)*	*Corresponding Questionnaire Question (Text)*		*value label*	*variable label*
EMS	**ems_0**	Early morning stiffness (hh:mm)	**Baseline**	How long is the patients early morning stiffness? (Must have been present over the last 6 weeks) in hh:mm		Early Morning Stiffness (hhmm)	
EMSF	**emsf_0**	Early morning stiffness fraction	**Baseline**	Derived		EMS Fraction	
TJC	**tjc_0**	Tender Joint Count	**Baseline**	Total No. of Tender Joints:		Total number of tender joints	
SJC	**sjc_0**	Swollen joint count	**Baseline**	Total No. of Swollen Joints		Swollen joint count	
SYM	**symmetric_0**	Symmetrical disease	**Baseline**	Symmetrical Disease?	SYM	symmetry	
NODULES	**nodules_0**	Rheumatic Nodules	**Baseline**	Any Nodules?	NODULES	rheumatoid nodules	
SWELLHW	**swellhw_0**	Swelling of hands or wrist	**Baseline**	Swelling of hands &/ or wrists?	SWELLHW	swelling hand/wrist	
HAQ	**haq_0**	Health Assessment Questionnaire	**Baseline**	Health Assessment Questionnaire (collated)		haq	
EROSION	**erosion_0**	Visible joint erosions from hands and feet X-rays	**Baseline**	Are any erosions visible?		erosions	
ANTICCP	**anticcp_0**	Anti CCP	**Baseline**	Not found	ANTICCP	anti-CCP	
ESR	**esr_0**	Erythrocyte Sedimentation Rate (mm/hr)	**Baseline**	ESR mm/hr		ESR	
DASESR	**das_esr_0**	Disease Activity Score ESR	**Baseline**	Derived		DASESR	
CRP	**crp_0**	C Reactive Protein (mg/dl)	**Baseline**	If not available, CRP:		CRP	
DASCRP	**das_crp_0**	Disease Activity Score CRP	**Baseline**	Derived		DASCRP	
HB	**hb_0**	Haemoglobin concentration (g/dl)	**Baseline**	Hb:		HB	
RAF	**raf_0**	Rheumatoid Factor	**Baseline**	Serum RA Factor?	RAF	RA Factor	
ORTHOEP	**ortho_ep_0**	orthopaedic Episode	**Baseline**	Not Found	ORTHOEP	orthopaedic episode 1	
COMORB	**comorbidity_0**	Presence of comorbidities	**Baseline**	Derived	COMORB	comorbidity-Baseline	
OHD	**ohd_0**	Organic Heart Disease	**Baseline**		OHD	Organic Heart Disease	

variable name	*New variable name*	*Description*	*Corresponding Questionnaire or follow up)*	*(baseline*	*Corresponding Questionnaire Question (Text)*	*value label*	*variable label*
IHD	**ihd_0**	Ischemic Heart Disease	**Baseline**		IHD	Ischemic Heart Disease	
PCAP	**pcap_0**	Primary Arrhythmias (PCAP)	**Baseline**		PCAP	Primary Arrhythmias (PCAP)	
CHF	**chf_0**	Congestive Heart Failure	**Baseline**		CHF	Congestive Heart Failure	
HYP	**hyp_0**	Hypertension	**Baseline**		HYP	Hypertension	
CVA	**cva_0**	Cerebral Vascular Accident	**Baseline**		CVA	Cerebral Vascular Accident	
PVD	**pvd_0**	Peripheral Vascular Disease	**Baseline**		PVD	Peripheral Vascular Disease	
DM	**dm_0**	Diabetes Mellitus	**Baseline**		DM	Diabetes Mellitus	
DETDM	**detdm_0**	Details DM	**Baseline**			Details DM	
DVT	**dvt_0**	DVT	**Baseline**		DVT	DVT	
RESP	**resp_0**	Respiratory Problems	**Baseline**		RESP	Respiratory Problems	
DEP	**dep_0**	depression	**Baseline**		DEP	depression	
ANX	**anx_0**	anxiety	**Baseline**		ANX	anxiety	
MAL	**mal_0**	Malignancies	**Baseline**		MAL	Malignancies	
HPD	**hpd_0**	Hepatobiliary Disease	**Baseline**		HPD	Hepatobiliary Disease	
REN	**ren_0**	Renal Disease	**Baseline**		REN	Renal Disease	
GID	**gid_0**	Gastrointestinal Disease	**Baseline**		GID	Gastrointestinal Disease	
HLIP	**hlip_0**	Hyperlipidaemia	**Baseline**		HLIP	Hyperlipidaemia	
OCOMORB1		Other Comorb 1				Other Comorb 1	
VISIT2	**visit_0**	visit 3-6month (x)				visit 3-6month (x)	

11.3.2 KPIC data access request form

KNEE PAIN IN THE COMMUNITY **Data Access Request Form**

Rheumatology, Orthopaedics and Dermatology, University of Nottingham

Instructions:
Applicants and collaborators requesting use of existing KPIC data are expected to provide information within each of the sections of this form. **Please submit fully completed form and signed agreement via email to Ana Valdes (ana.valdes@nottingham.ac.uk) and annotate "Data Access Request" in the subject line. Incomplete and/ or UNSIGNED forms will delay the request process.**

Section 1 - Data Applicant Details & Institutional Information		
1.1	**Applicant Details**	Title: Dr Position: PhD Student Name: Onosi Ifesemen Email: mbxosi@nottingham.ac.uk Telephone Number:
1.2	**Co Applicant/ Supervisor Details**	Title: Prof Position: Prof Rheumatology/supervisor Name: David Andrew Walsh Email: david.walsh@nottingham.ac.uk Telephone Number: 0115 823 1766 Title: Dr Position: Research fellow/supervisor Name: Dan McWilliams Email: dan.mcwilliams@nottingham.ac.uk Telephone Number: 0115 82 31942 Title: Prof Position: Professor in Musculoskeletal Genetics Name: Ana M Valdes Email: ana.valdes@nottingham.ac.uk Telephone Number:
1.3	**Institution Details**	Name: University of Nottingham and Sherwood Forest Hospitals NHS Foundation Trust Department: Academic Rheumatology/Rheumatology Postal Address: CSB, City Hospital, Hucknall Road, Nottingham, NG5 1PB
1.4	**Date of Request**	24 July 2020
1.5	**Have you requested data from the KPIC database before?**	Yes – Please provide project title of previous request: Walsh: Mediators of the predictive effect of self-report 'central

		mechanisms' on persistent knee pain in the Knee Pain in the Community (KPIC) cohort Walsh: Influence of peripheral and central pain mechanisms on quadriceps strength in people with knee pain.

Section 2 - Project Details		
2.1	**Project Title**	Fatigue in the Community
2.2	**Proposed Start Date**	24 July 2020
2.3	**Proposed End Date**	31 July 2021

Section 3 - Ethical Approval		
3.1	**Does the study supporting your project have ethical approval from an Ethics Committee or an Institutional Review Board?**	Not Applicable (You may tick this box if you intend to carry out analysis of existing data where the KPIC ethics approval will operate)
3.2	**Please specify arrangements for obtaining appropriate ethical approval**	NA

Section 4 - Scientific Project Details

Questionnaire Data

Please indicate details on the traits/data required

Baseline Questionnaire	Year 1 Questionnaire
Age, sex, Weight, Height, BMI, Occupation, Medication, Comorbidities: Q1 – 6, Q9	New comorbidities:Q2 Medication: Q7
Pain: painDETECT, ICOAP	painDETECT, ICOAP
Pain: Section 6, Q34 – Q38	Section 4
Mental Health: Section 7: Q39	Section 5
QoL: Section 8	Section 6 & 7
Others	
GPPAQ	GPPAQ
Central mechanisms Traits – all individual components & total	Central mechanisms Traits – all individual components & total

		score / score Biomarkers: CRP / Biomarkers: CRP Central sensitization: QST / Central sensitization: QST Gait Assessment – Gaitrite & RSscan data / Gait Assessment – Gaitrite & RSscan data Muscle strength / Muscle strength *GPPAG (General Practice Physical Activity Questionnaire), QST (Quantative Sensory Testing), CRP (C- Reactive Protein), QoL (Quality of Life), ICOAP (Intermittent and Constant Osteoarthritis Pain)
	Please indicate participants required	Total population

Section 5- Scientific Outline

Please provide a 1-2 page outline of your proposal highlighting the specific project requirements for the KPIC data specified above stating the rationale for using this data, including other study methods considered. Please be as precise as possible with regards to trait data required listing the questionnaire areas of interest and inclusion/exclusion criteria of any participants to be studied or approached for possible recruitment into a new study.

Fatigue in the community: Burden and Risk Factors

Background:

Fatigue can be defined as a subjective experience of intense tiredness and exhaustion, and not relieved by rest. It is one of the commonest symptoms reported in primary care and is associated with poor long-term quality of life. Fatigue has a profound effect on the work ability of people, and it is estimated to cost employers about $136.4 billion/year in the USA alone(1).

Fatigue is a cardinal symptom of many chronic diseases such as certain cancers and chronic inflammatory conditions e.g rheumatoid arthritis however fatigue has been known to persist, even in the absence of underlying disease. The exact cause of fatigue is not known, although several causal mechanisms including organic, psychosocial, and environmental factors have been proposed. These mechanisms are often related and the resulting fatigue, a consequence of a combination of several factors (2).

One of such mechanisms is central sensitization (CS). Central sensitization is characterized by amplified painful responses and hypersensitivity to environmental stimuli. It is caused by amplification of sensory input across multiple bodily systems. CS is present, to some degree in many inflammatory and non-inflammatory MSK disorders associated with fatigue e.g fibromyalgia and rheumatoid arthritis (RA)(3).

Several studies have examined risk factors associated with fatigue in many chronic diseases and in the general population. Factors such as mental health, sleep, socioeconomic status, and physical activity tend to be identified as risk factors for fatigue in the general population, while inflammation, pain and depression are risk factors identified in studies chronic disease

populations (4-6).

Rationale: Few studies have examined the risk factors for fatigue accounting for the contribution of inflammation, pain and central mechanisms in the general population.

Aim: Investigate the prevalence and risk factors for fatigue in the general population.

Objectives:

- Describe the prevalence of fatigue in the community
- Identify risk factors associated with concurrent fatigue at baseline and year 1
- Investigate baseline factors associated with incident fatigue in year 1
- Examine the relationship between CS and fatigue in the community

Hypothesis: Central sensitization is associated with fatigue independent of other risk factors.

Methods:

Population: Baseline and year 1 data from all KPIC participants.

Primary outcome measure:

Fatigue: Obtained from questionnaire data: Q55" In the past month, did you feel tired on most days?"; This will be dichotomized based on responses. Participants who ticked 'often' and 'always' will be classed as the fatigue group(7).

Explanatory variables:

Central sensitization:

Quantitative Sensory Testing - Pressure pain detection threshold (PPT) and temporal summation (TS)

Covariates: (details in 'data/traits required' the section above)

Demographics: Age, sex, BMI, Occupation, Medication, Comorbidities

Sleep

Quality of Life

Mental Health: HADS-anxiety, HADS-depression,

Pain: Pain Catastrophizing Scale, painDETECT, ICOAP

Central mechanisms trait score - anxiety, depression, catastrophizing, neuropathic-like pain, fatigue, sleep disturbance, pain distribution, and cognitive impact

Widespread pain from mannikin, ACR Fibromyalgia criteria

Physical Activity: General Practice Physical Activity Questionnaire,

Inflammation: CRP

Frailty: Fried frailty phenotype(8)

Statistical analysis:

Descriptive statistics of the study population will be provided at baseline and in year 1.

Prevalence of fatigue will be calculated at baseline standardized to age and sex with the European standard population 2013.

Concurrent risk factors for fatigue at baseline/year 1 fatigue, and baseline risk factors for year 1 fatigue will be assessed using logistic regression

Causal theory will be used to assess the relationship between CS and fatigue.

Sensitivity analyses:

Analysis will be conducted on subpopulations with and without chronic pain. Other sensitivity analyses will be based on the results of the primary analysis to interrogate the validity of our results.

References

1. Ricci JA, Chee E, Lorandeau AL, Berger J. Fatigue in the US workforce: prevalence and implications for lost productive work time. Journal of occupational and environmental medicine. 2007;49(1):1-10.
2. Hewlett S, Chalder T, Choy E, Cramp F, Davis B, Dures E, et al. Fatigue in rheumatoid arthritis: time for a conceptual model. Rheumatology. 2010;50(6):1004-6.
3. Latremoliere A, Woolf CJ. Central sensitization: a generator of pain hypersensitivity by central neural plasticity. J Pain. 2009;10(9):895-926.
4. Engberg I, Segerstedt J, Waller G, Wennberg P, Eliasson M. Fatigue in the general population- associations to age, sex, socioeconomic status, physical activity, sitting time and self-rated health: the northern Sweden MONICA study 2014. BMC Public Health. 2017;17(1):654-.
5. Galland-Decker C, Marques-Vidal P, Vollenweider P. Prevalence and factors associated with fatigue in the Lausanne middle-aged population: a population-based, cross-sectional survey. BMJ open. 2019;9(8):e027070.
6. Beiske AG, Loge JH, Hjermstad MJ, Svensson PE. Fatigue in Parkinson's disease: Prevalence and associated factors. Movement Disorders. 2010;25(14):2456-60.
7. Morley JE, Malmstrom TK, Miller DK. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. J Nutr Health Aging. 2012;16(7):601-8.
8. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in Older Adults: Evidence for a Phenotype. The Journals of Gerontology: Series A. 2001;56(3):M146-M57.

11.3.3 Jyvaskyla Early Rheumatoid Arthritis Data Request

1. Lead Applicant

Name	Dr Onosi Ifesemen
Position	PhD Student
Department/institution address	Arthritis Research UK Pain Centre, NIHR Biomedical Research Centre, Division of ROD, University of Nottingham
Email	Onosi.ifesemen@nottingham.ac.uk

2. Co-applicants

Name	Dr Daniel McWilliams
Position	Post-Doctoral Research Fellow
Department/institution address	Arthritis Research UK Pain Centre, NIHR Biomedical Research Centre, Division of ROD, University of Nottingham
Email	Dan.mcwilliams@nottingham.ac.uk

Name	Dr Elena Nikiphorou
Position	Rheumatologist
Department/institution address	Dept of Rheumatology, KCL, London
Email	enikiphorou@gmail.com

Name	Professor David A. Walsh
Position	Professor of Rheumatology
Department/institution address	Arthritis Research UK Pain Centre, NIHR Biomedical Research Centre, Division of ROD, University of Nottingham
Email	mailto:david.walsh@nottingham.ac.uk

3. Project

Title	Understanding Pain and fatigue in early Rheumatoid Arthritis
Source of funding	Early Rheumatoid Arthritis Network (ERAN)
Start date	01/01/2018
Duration	36 months
Collaborators	None

4. Ethical Approval

Is ethical approval required for this project?	NO
IF NO, reasons why not	Analysis of already collected data

5. Information Governance

We only require linked-anonymized data; with each individual denoted by a unique anonymous code that is supplied by the Jyvaskyla cohort managers. These data will be stored in the desktop computers within Academic Rheumatology and secure servers of the University of Nottingham (this means it is not on OneDrive). Data will be stored in accordance with the GDPR guidelines and will not be shared with anyone outside the team of collaborators without prior written approval from the Jyvaskyla cohort managers. All data analysis will be carried out in the University of Nottingham.

At the end of the project, data will be archived in the University of Nottingham for audit purposes. The data will not be used for other analyses and will not be merged with other datasets. Data will be deleted after a period of 10 years post Onosi Ifesemen's PhD thesis submission.

Results of data analysis will be presented in a PhD thesis and academic publications.

6. Data Protection

This project does not involve linking to patient identifiable data from other sources.

7. Project Outline and Analysis protocol

Introduction/Background

Fatigue and pain are prevalent in people with Rheumatoid Arthritis, and persist despite breakthroughs in pharmacological management of the disease. Pain and fatigue are thought to be caused by overlapping mechanisms however much is still unknown.

Several models have been proposed to explain the relationships among pain, fatigue, and inflammation in RA however these models are largely based on cross-sectional data, with few longitudinal studies and fewer studies based on people with early disease. Furthermore, these models portray fatigue and pain as consequences of inflammation but this relationship does not explain the persistence of pain and fatigue in RA patients with low levels of inflammatory markers and low disease activity scores.

Preliminary Results in ERAN already show a significant association between pain and fatigue independent of inflammation.

Hypothesis

Our hypothesis is that pain may be responsible for fatigue experienced in these individuals and that this occurs independently of inflammatory activity.

Method

The analyses would be conducted using data from two RA cohorts, the Early Rheumatoid Arthritis Network inception cohort, and the Jyvaskyla Cohort. These cohorts represent rich data sources containing information on individuals with RA from the time of diagnosis and over time. This allows for the examination of disease progression as well as the progression of our primary outcomes. Our selection of covariates is based on a literature search and a proposed conceptual model . Two datasets were selected in order to establish the consistency of research findings.

Objectives

1. Describe the longitudinal relationship between fatigue and pain in Rheumatoid Arthritis. Investigate whether this is explained by inflammation.

   1a. Subgroup individuals according to their fatigue trajectories.

2. Examine the temporality of pain and fatigue in RA

Statistical Analysis

1. Baseline data will be examined using simple descriptive statistics and regression analysis appropriate to the outcome measures of pain and fatigue.

2. Longitudinal relationships will be explored with growth curve modelling and group based trajectory modelling, using the structural equation modelling framework. Baseline predictors of fatigue progression will be investigated.

3. Temporality between pain and fatigue will be assessed with cross lagged regression analysis also within the structural equation modelling framework. Time varying predictors of fatigue progression will be investigated.

8. Data Requested

Data requested (if available):

- Demographic Data (age, sex, socioeconomic status, smoking status, occupation, education, BMI)
- Longitudinal clinical/laboratory and patient reported outcome measures. Baseline and all follow up time points.
    - Fatigue
    - Pain
    - Measures of disability - Health Assessment Questionnaire (HAQ)
    - Measures of disease activity/ Inflammation - Swollen Joint Count (SJC), Tender Joint Count (TJC), Erythrocyte Sedimentation rate (ESR), C - Reactive Protein (CRP), Global Assessment of Disease Activity, Serological Data (Rheumatoid Factor (RF), Anti-Citrullinated Protein (CCP)
    - Sleep
    - Medication use (Disease Modifying Anti-Rheumatic Drugs (DMARDs), Biologics, Glucocorticoids. Anti-Inflammatory medication, prescribed analgesia
    - Co-morbidities – Composite measures or presence or absence of individual co-morbidities
    - Composite Disease Activity Scores (if available)
    - Anxiety , Depression, Mental health
- Longitudinal outcomes (e.g death, surgery)[2]

At end of the study, we should produce 2 – 3 academic publications describing the longitudinal relationship between pain and fatigue, Identifying clusters of individuals with different fatigue progressions over time (phenotyping) and describing the temporality of pain and fatigue in RA.

[2] Please note that outcome variables are requested to help us understand the cause of missing data (where applicable).

11.3.4 Jyvaskyla ethics committee approval

ksshp
Keski-Suomen sairaanhoitopiiri

KESKI-SUOMEN SAIRAANHOITOPIIRI

TIETEELLISEN TUTKIMUKSEN LUPAHAKEMUS
(sis. potilastietojen luovutus- ja käyttöhakemuksen)

Allekirjoitettu lomake lähetetään postissa/sisäpostissa tutkimuskoordinaattorille Liitteet lähetetään sähköpostilla. Jos tutkimuksessa ei tarvita potilastietoja, niin päätutkijan allekirjoitus riittää; muuten kaikki potilastietoja käsittelevät allekirjoittavat päätutkijan lisäksi. Tutkijan tulee sopia yksikön palvelujohtajan ja palveluesimiehen kanssa ennen luvan hakemista tutkimuksen toteutuksesta ko. yksikössä.

1 Päätutkijaa koskevat tiedot

Suku- ja etunimi		Oppiarvo
Sokka-Isler Tuulikki		Prof
Nykyinen työnantaja KSSHP ☒ Kyllä ☐ Ei, mikä?		Nykyinen virka/toimi Prof, OyI
Puhelin toimeen 040 735 2087	Puhelin kotiin sama	Sähköpostiosoite tuulikki.sokka-isler@ksshp.fi

2 Tutkimusta koskevat tiedot

Tutkimuksen nimi/lyhenne
Fatigue in Rheumatoid Arthritis
Uupumus/väsymys nivelreumassa
(FATIGUE)

Tutkimuksen suorituspaikka/t (yksikkö/t)
KSSHP, reumatutkimus

Yksikön/yksiköiden palvelujohtajan ja palveluesimiehen kanssa on sovittu tutkimuksen toteutuksesta
☒ Hyväksyntä, kenen/keiden
Jouni Ranua

Tutkimuksen vastuuhenkilö, TVH (tarv. yhteystiedot), jos muu kuin hakija
Hakija

Liittyykö tutkimus muuhun meneillään olevaan tutkimukseen/laajempaan tutkimusohjelmaan
☒ Ei
☐ Kyllä, mihin?

Tutkimus tehdään useassa eri toimipisteessä	Tutkimus on
☒ Ei ☐ Kyllä ☐ Kansallinen ☐ Kansainvälinen	☒ Uusi tutkimus ☐ Jatkotutkimus
Tutkimuksen aloitus pvm Heti kun lupa saatu	Tutkimuksen arvioitu päättymis pvm 31.12.2030

Tutkimus on
☒ Lääketieteellinen
☐ Hoitotieteellinen
☐ Muu terveystieteellinen, mikä?
☐ Muu, mikä?

Tutkija tarvitsee luvan potilastietojen käyttöön potilastietojärjestelmistä/arkistosta
☐ Ei
☒ Kyllä (Tietosuojaseloste ja Tietoturvariskien itsearviointi tutkimuksessa pakolliset)

8. Liitteet (toimitetaan sähköpostilla; *merkityt pakollisia)

- ■ Tutkimussuunnitelma (sis. tiivistelmän)*
- ■ Tietosuojaseloste (Rekisterinpitäjän organisaation ohjeistuksen mukainen)*
- ■ Tietoturvariskien itsearviointi tutkimuksessa (Rekisterinpitäjän organisaation ohjeistuksen mukainen)*
- ☐ THL:n / STM:n lupa
- ☐ Tutkittavien suostumukset
- ☐ Muut asiakirjat, mitkä

9. Sitoumukset (sitoumuksen allekirjoittavat kaikki henkilöt, jotka tutkimusta tehtäessä käsittelevät salassa pidettäviä tietoja)*

Sitoudun noudattamaan EU – asetuksia, Suomen lakeja, asetuksia ja säädöksiä tieteellisestä tutkimustoiminnasta sekä Keski-Suomen sairaanhoitopiirin tieteellisestä tutkimuksesta ja sen raportoinnista antamia ohjeita ja sääntöjä.

Sitoudun siihen, etten käytä saamiani tietoja potilaan/asiakkaan tai hänen läheisensä vahingoksi tai halventamiseksi taikka sellaisten muiden etujen loukkaamiseksi, joiden suojaksi on säädetty salassapitovelvollisuus enkä luovuta saamiani tietoja muille kuin tutkimus/esitysluvassa mainituille henkilöille.

Päivämäärä Allekirjoitus 06/01/2020 Nimenselvennys Onosi Ifesemen	Organisaatio Pain Centre Versus Arthritis, Academic Rheumatology, University of
Päivämäärä Allekirjoitus _____ Nimenselvennys	Organisaatio
Päivämäärä Allekirjoitus _____ Nimenselvennys	Organisaatio
Päivämäärä Allekirjoitus _____ Nimenselvennys	Organisaatio
Päivämäärä Allekirjoitus _____ Nimenselvennys	Organisaatio

Sitoumukset (sitoumuksen allekirjoittavat kaikki henkilöt, jotka tutkimusta tehtäessä käsittelevät salassa pidettävä tietoja)

9.12.2019	Tuulikki Sokka-Isler	Tuulikki Sokka-Isler
Pvm.	Allekirjoitus ja nimenselvennys	
9.12.2019	Kirsi Paalanen	Kirsi Paalanen
Pvm.	Allekirjoitus ja nimenselvennys	
9.12.2019	Seija Salonen	Seija Salonen
Pvm.	Allekirjoitus ja nimenselvennys	
Pvm.	Allekirjoitus ja nimenselvennys	
Pvm.	Allekirjoitus ja nimenselvennys	
Pvm.	Allekirjoitus ja nimenselvennys	
Pvm.	Allekirjoitus ja nimenselvennys	
Pvm.	Allekirjoitus ja nimenselvennys	
Pvm.	Allekirjoitus ja nimenselvennys	

PÄÄTÖS

☒ Tutkimuslupa myönnetään hakemuksen mukaisena

- Myönnetään potilastietojärjestelmän käyttöoikeus (tutkijatunnukset) sitoumuksen allekirjoittaneille omille tutkijoille. Tunnukset haetaan tutkimuskoordinaattorilta tutkijatunnuksien hakulomakkeella
- Ulkopuolisille tutkijoille aineistot pääsääntöisesti haetaan sopimuksen mukaan (maksullinen)

☐ Myönnetään edellyttäen, että _____

☐ Hakemus hylätään seuraavin perustein _____

Päivämäärä	Tutkimusluvan myöntäjän nimi, arvo ja nimen selvennys
24.01.2020	Vesa Kataja johtajaylilääkäri Keski-Suomen sairaanhoitopiiri

KSSHP, tutkimuslupa/p1

Liitteet (toimitetaan sähköpostilla)

Lääketieteellinen ja muu tutkimus

- ▣ Tutkimussuunnitelma
  (sis. tutkimuskeskukset ja tutkijat, tiivistelmän, kustannusarvion ja rahoitussuunnitelman)
- ☐ Eettisen toimikunnan lausuntokopio (tutkimuslain alaiset; yliopiston eettinen)
- ▣ Tietosuojaseloste (tutkimusrekisteri) *
- ▣ Tietoturvariskien itsearviointi tutkimuksessa

(Rekisterinpitäjän organisaation ohjeiden mukainen)*

- ☐ Biopankkitiedote ja – suostumus
- ☐ Muut lupakopiot
  (STM:n / THL:n lupa, Fimea, Valvira tmv.)
- ☐ Muu materiaali tarv., jos ei eettisen toimikunnan lausuntoa
  (mm .tiedote/saate, suostumuslomake, kyselyt, haastattelulomake, muut mittarit)

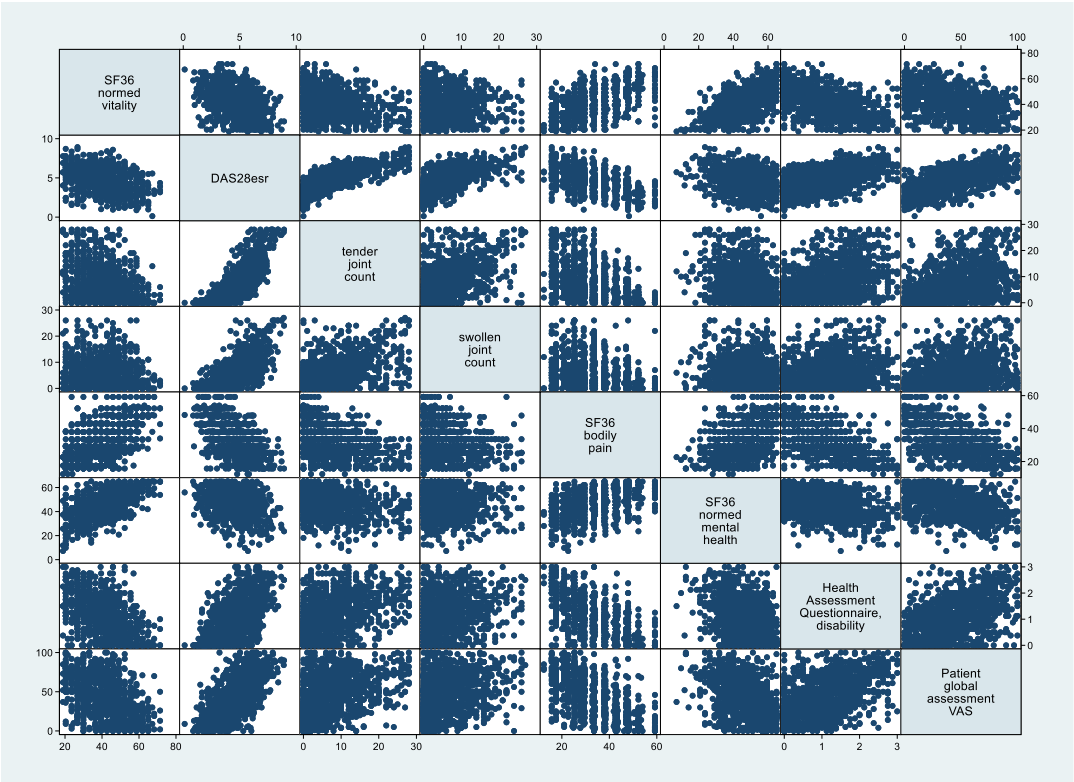
11.4 Cross-sectional study of fatigue in RA Appendices

Appendix Table 11.1:European Standard Population (ESP)

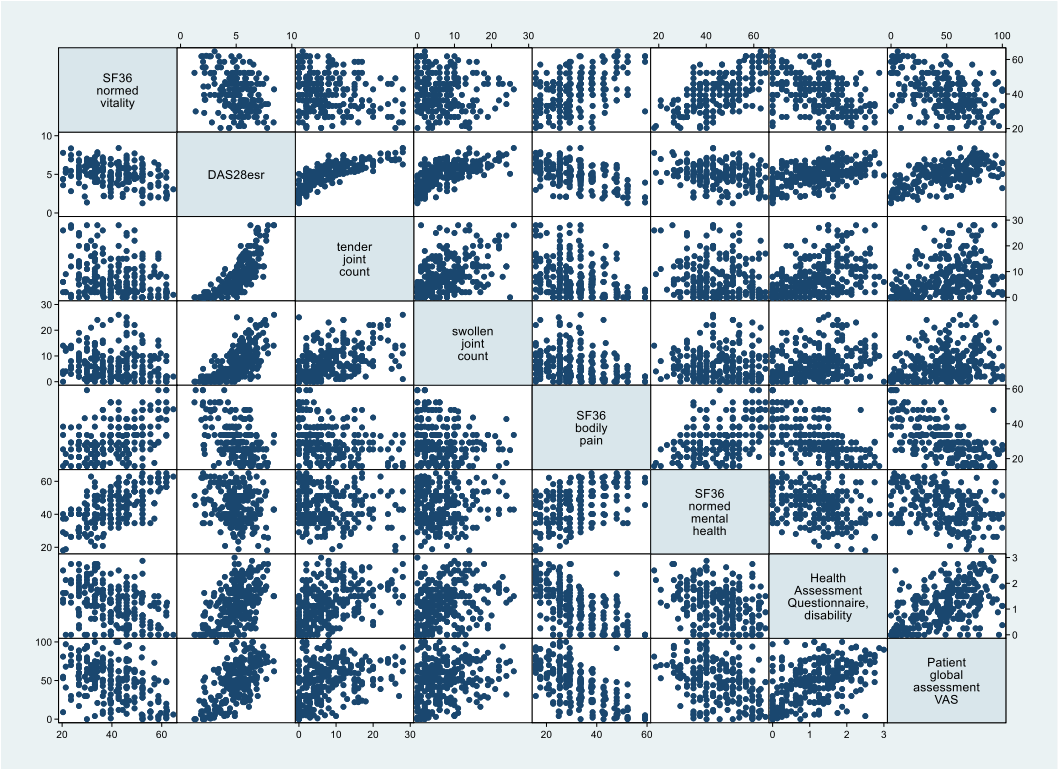
Age Group	Sex	ESP
0-4 years	Male	5000
0-4 years	Female	5000
5-9 years	Male	5500
5-9 years	Female	5500
10-14 years	Male	5500
10-14 years	Female	5500
15-19 years	Male	5500
15-19 years	Female	5500
20-24 years	Male	6000
20-24 years	Female	6000
25-29 years	Male	6000
25-29 years	Female	6000
30-34 years	Male	6500
30-34 years	Female	6500
35-39 years	Male	7000
35-39 years	Female	7000
40-44 years	Male	7000
40-44 years	Female	7000
45-49 years	Male	7000
45-49 years	Female	7000
50-54 years	Male	7000
50-54 years	Female	7000
55-59 years	Male	6500
55-59 years	Female	6500
60-64 years	Male	6000
60-64 years	Female	6000
65-69 years	Male	5500
65-69 years	Female	5500
70-74 years	Male	5000
70-74 years	Female	5000
75-79 years	Male	4000
75-79 years	Female	4000
80-84 years	Male	2500
80-84 years	Female	2500
85-89 years	Male	1500
85-89 years	Female	1500
90plus years	Male	1000
90plus years	Female	1000
Source:https://www.opendata.nhs.scot/dataset/standardpopulations/resource/29ce4cda-a831-40f4-af24-636196e05c1a		

Appendix Table 11.4 shows the population values for the European standard population for all age groups by age and sex. These values were aggregated to match ERAN population.

Scatter plot matrices

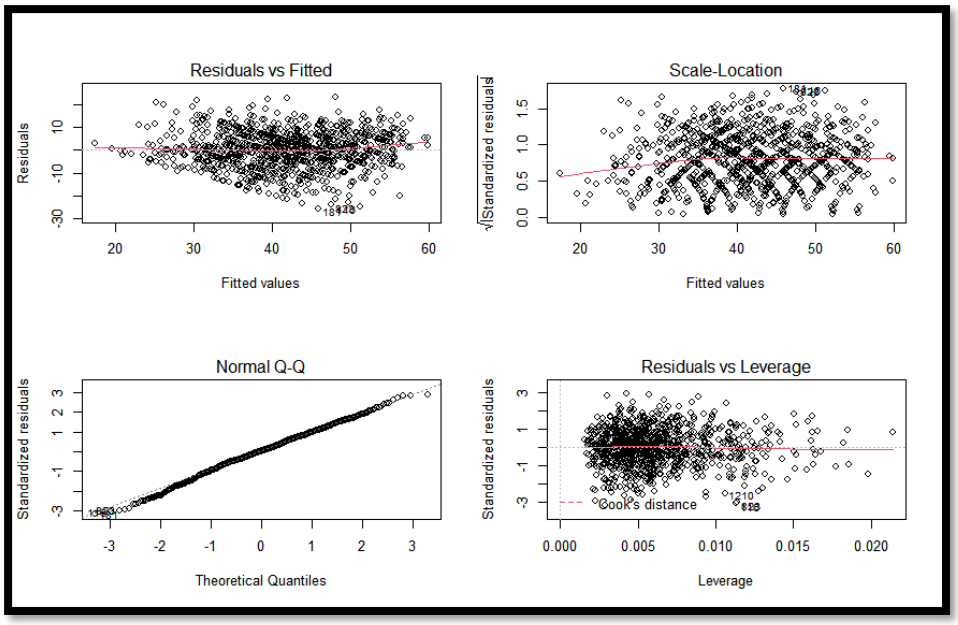


Appendix figure 11.1: Scatterplot matrix of covariates in the total ERAN population at baseline

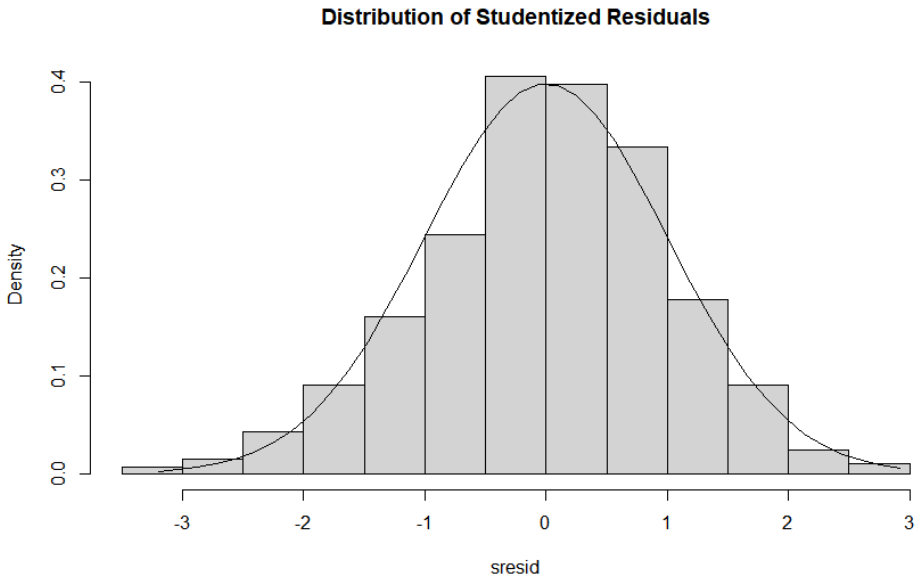


Appendix Figure 11.2: Scatter plot matrix of covariates in the ERAN population with a disease duration of less than 6 months

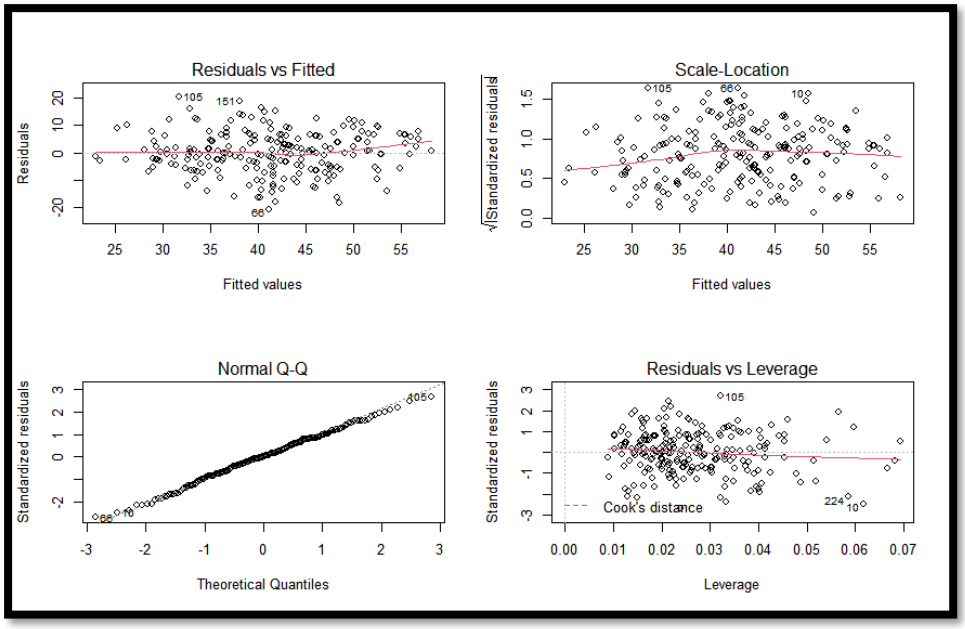
Regression Linear Regression Diagnostic plots



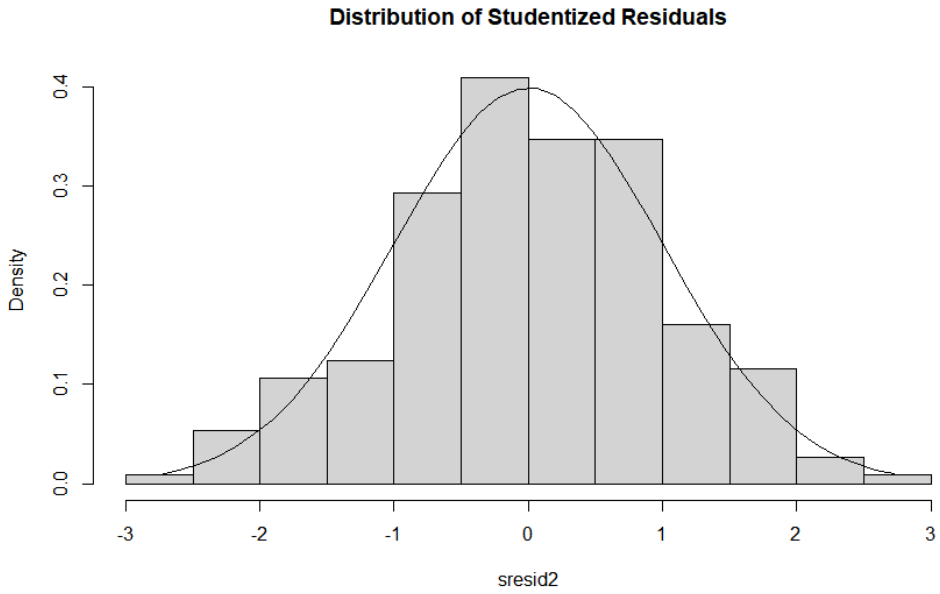
Appendix Figure 11.3:Graphical presentation of residuals from the total ERAN population.



Appendix Figure 11.4:Histogram of residuals from the multivariable model of the total ERAN population



Appendix Figure 11.5:Graphical examination of residuals from the multiple regression model from the ERAN sub-population



Appendix Figure 11.6:Histogram of residuals from ERAN participants with a disease duration of less than 6 months

Appendix Table 11.2: Test for multicollinearity Total population

Variable	VIF	Tolerance
Age	1.08	0.943598
Sex	1.10	0.893142
Pain	1.96	0.534537
TJC	1.36	0.733317
Mental Health	1.39	0.751424
Haq	1.94	0.597938
Mean VIF	1.47	
VIF- Variance Inflation Factor		

Appendix Table 11.3: Test for multicollinearity population less than 6 months duration

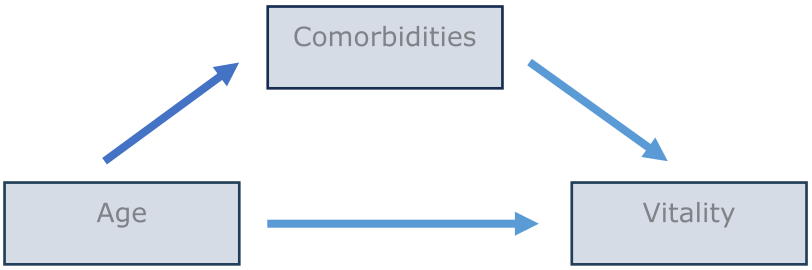
Variable	VIF	Tolerance
Age	1.06	0.943598
Sex	1.12	0.893142
Pain	1.87	0.534537
Mental Health	1.33	0.751424
Haq	1.67	0.597938
Mean VIF	1.41	
VIF- Variance Inflation Factor		

11.4.1 Explaining the mediation effect of comorbidities on the association between age and fatigue (ERAN data)

The proposed effect of comorbidities on the association between age and fatigue was tested using mediation analysis conducted according to the procedure described by Baron and Kenny(Baron and Kenny, 1986). Estimates were produced using the Imai et al algorithm, and confidence intervals using quasi-Bayesian Monte Carlo approximation. Mediation was conducted using the 'mediate' package on R statistical software version 4.2.1.

The figure below represents the assumed mediation model.

Findings, represented in table 11.3.1 below showed a significant small mediational effect of comorbidities on the association between age and fatigue in the negative direction, -0.05 (CI: -0.07 to -0.02, p<0.001). In the presence of the mediating variable (comorbidities), the association of age on fatigue was not statistically significant, 0.05 (CI: -0.001 to 0.10, p=0.07).



Appendix figure 211.6.1: Mediation Model (Assumed)

Appendix Table 11.3.1 Mediation Effect Estimates (Age of onset)

	Coefficient	Confidence Interval	P Value
Indirect Effect	-0.04531	-0.07 to -0.02	<0.001
Direct Effect	0.09555	0.04 to 0.15	<0.001
Total Effect	0.05024	-0.001 to 0.10	0.066

Values are unstandardized regression coefficients

11.5 Longitudinal course of fatigue in RA Appendices

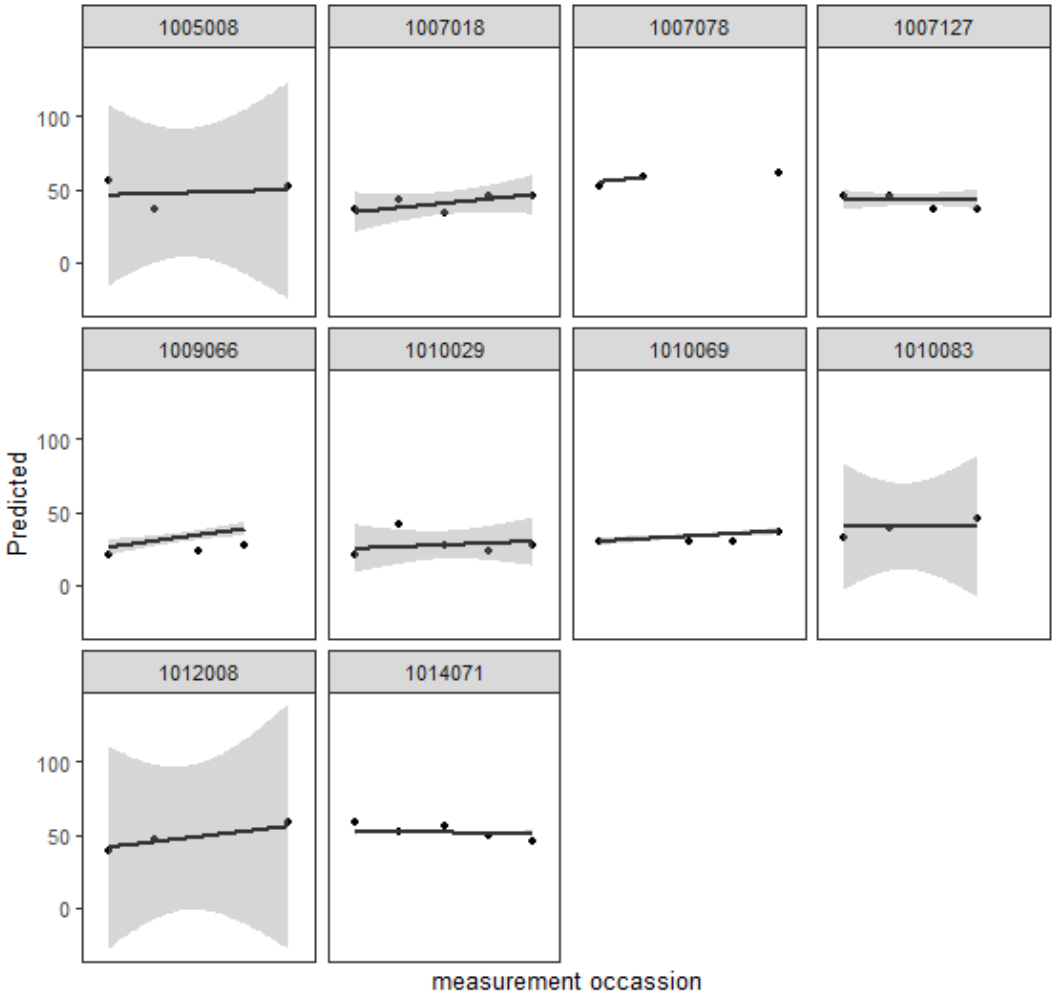
Appendix Table 11.4: GROLTS Checklist for LGMM

GROLTS Checklist		
No	Checklist Item	Reported ?
1	Is the metric of time used in the statistical model reported?	Yes
2	Is information presented about the mean and variance of time within a wave?	Yes
3a.	Is the missing data mechanism reported?	Yes
3b.	Is a description provided of what variables are related to attrition/missing data?	Yes
3c.	Is a description provided of how missing data in the analyses were dealt with?	Yes
4	Is information about the distribution of the observed variables included?	Yes
5	Is the software mentioned?	Yes
6a.	Are alternative specifications of within-class heterogeneity considered (e.g., LGCA vs. LGMM) and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	Yes
6b.	Are alternative specifications of the between-class differences in variance covariance matrix structure considered and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	Yes
7	Are alternative shape/functional forms of the trajectories described?	Yes
8	If covariates have been used, can analyses still be replicated?	Yes
9	Is information reported about the number of random start values and final iterations included?	Yes
10	Are the model comparison (and selection) tools described from a statistical perspective?	Yes
11	Are the total number of fitted models reported, including a one-class solution?	Yes
12	Are the number of cases per class reported for each model (absolute sample size, or proportion)?	Yes
13	If classification of cases in a trajectory is the goal, is entropy reported?	Yes
14a.	Is a plot included with the estimated mean trajectories of the final solution?	Yes
14b.	Are plots included with the estimated mean trajectories for each model?	Yes
14c.	Is a plot included of the combination of estimated means of the final model and the observed individual trajectories split out for each latent class?	Yes
15	Are characteristics of the final class solution numerically described (i.e., means, SD/SE, n, CI, etc.)?	Yes
16	Are the syntax files available (either in the appendix, supplementary materials, or from the authors)?	Not yet

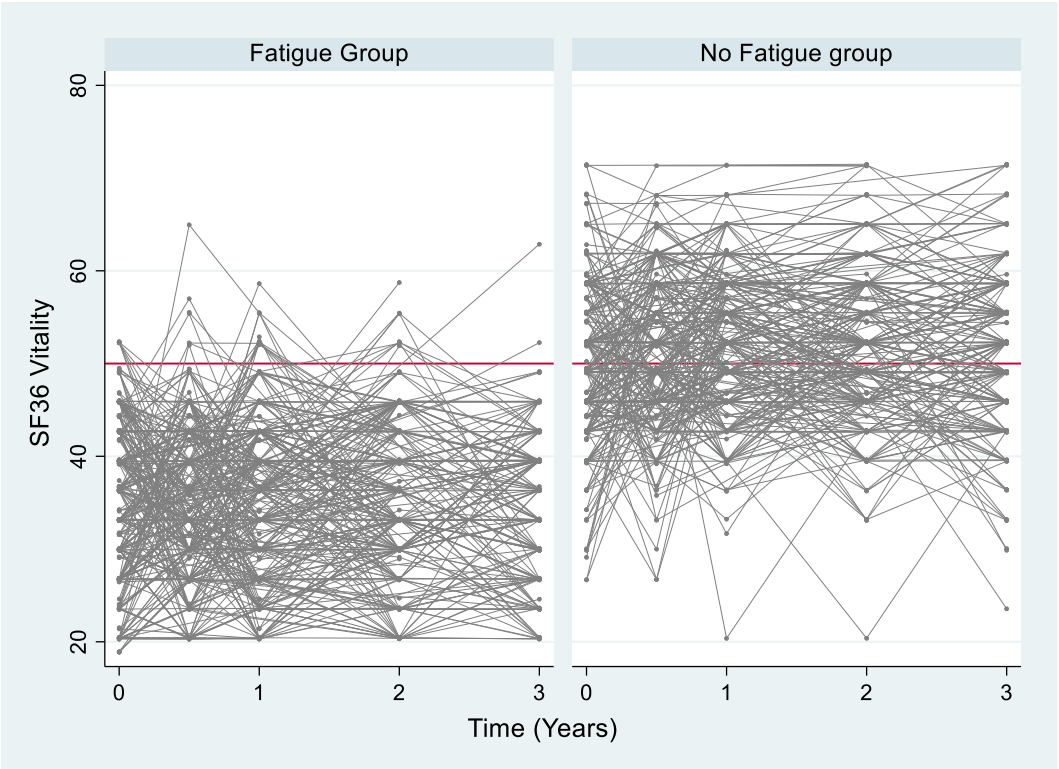
Appendix Table 11.5:descriptives characteristics of the ERAN population at each measurement occasion

ERAN	**N=1236**	**n=1192**	**n=1179**	**n=1151**	**n=1138**
Characteristics	Baseline	6mths	yr1	yr2	Yr3
SF36VT	41.82 (11.14)	42.85 (11.61)	43.78(12.17)	43.34(11.80)	43.83(11.62)
Age	57.01 (13.80)	57.33 (13.80)	58.01(13.59)	58.44(13.73)	59.11(13.72)
HAQ	1.08 (0.76)	1.00 (0.78)	0.96 (0.79)	0.97 (0.80)	0.95 (0.80)
PGA	43.50 (25.61)	39.07 (25.70)	37.85 (26.92)	36.29 (26.21)	35.50 (26.01)
Hb	13.1 (1.39)	13.20 (1.39)	13.2 (1.37)	13.17 (1.40)	13.25 (1.37)
DAS28ESR	4.68 (1.56)	3.95 (1.56)	3.80 (1.62)	3.65 (1.67)	3.44(1.64)
SF36BP	33.77 (10.62)	36.15 (10.41)	37.38 (11.32)	37.63 (10.75)	38.54 (11.04)
SF36MH	46.65 (37.4 - 53.9)	48.39 (39.70-56.66)	48.39 (40.13 - 56.66)	51.15 (40.13 - 56.66)	51.15 (40.13 - 56.66)
RDCI	1 (0 - 2)	1 (0 - 2)	1 (0 - 2)	1 (0 - 2)	1 (0 - 2)
Disease duration	12 (7 - 21)	16 (12 - 25)	24 (19 - 33)	36 (31 - 46)	49 (44 - 58)
TJC	5 (2- 11)	3 (0 - 8)	2 (0 - 8)	2 (0 - 6)	1 (0 - 6)
SJC	4 (1- 9)	2 (0 - 6)	1 (0 - 5)	1 (0 - 4)	1 (0 - 3)
ESR	24 (12 - 41)	16 (9 - 30)	17 (8 - 30)	15 (8 - 29)	15 (8 - 27)
BMI	26.84 (23.92 - 30.46)	27.18 (24.25 - 30.79)	27.40 (24.2 - 31.20)	27.05 (24.24- 30.97)	27.54 (24.54 - 30.86)
Medication use*	935 (78.31)	820 (90.01)	763 (89.55)	669 (89.56)	591 (90.23)
Seropositivity	655 (61.04)	662 (60.46)	672 (60.05)	678 (59.89)	683 (59.70)
Never Smoker	477 (40.38)	357 (40.38)	328 (39.95)	283 (38.14)	238 (38.14)
Ex_Smoker	404 (34.50)	305 (36.94)	304 (36.94)	248 (39.74)	248 (39.74)
Current Smoker	311 (25.11)	222 (25.11)	191 (23.21)	1138 (22.12)	138 (22.12)
Erosions present	331(29.40)	335 (29.49)	355(30.68)	367 (31.42)	396 (33.67)
Nodules present	100 (83)	83 (7.89)	83 (7.89)	83 (7.89)	83 (7.89)
nSF36VT - Normed SF36 Vitality , Age - Age at visit, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Hb - haemoglobin Concentration, DAS28ESR - Disease Activity Score (ESR), nSF36BP - Normed SF36 Bodily Pain, nSF36MH - Normed SF36Mental health, RDCI - Rheumatic Disease Comorbidity Index, Duration - Disease Duration in Months, TJC - Tender Joint Count, SJC - Swollen Joint Count, ESR - Erythrocyte Sedimentation rate, BMI - Body Mass Index. * - DMARDS					

Appendix Figure 11.8:Predicted growth curves from 10 randomly selected study participants



Appendix Figure 11.9: Trajectories of Vitality by fatigue groups



Appendix Figure 11.8 shows fatigue trajectories by group allocation, over the study duration of 3 years. The graph shows that the fatigue group are most likely to have Vitality values below the normed UK average value of 50 (straight red line), and the no-fatigue group, values above the UK normed value.

Appendix Table 11.6: Fatigue progression over study duration using non-imputed data

		Values				
		No Growth	Linear Growth	Quadratic growth	spline1	spline
Model Parameters	Time	0.18 (0.04 to 0.32) *	0.17 (0.03 to 0.32) *	0.91 (0.45 to 1.36) *	1.17 (0.43 to 1.91) *	
	$Time^2$			-0.13 (-0.20 to -0.05) *		
	S1				-1.15 (-1.97 to -0.32) *	1.23 (0.51 to 1.95) *
	s2					0.01 (-0.22 to 0.19)
	Constant	42.53 (41.75 to 43.31) *	42.54 (41.78 to 43.31) *	42.06 (41.24 to 42.87) *	42.00 (41.17 to 42.83) *	42.01 (41.17 to 42.84) *
	Log Likelihood	-10539.51	-10529.21	-10526.10	-10514.67	-10515.18
	Akaike Inf. Crit.	21087.03	21070.43	21066.03	21049.34	21050.35
	Bayesian Inf. Crit.	21110.9	21106.23	21107.80	21109.02	21110.03

*p<0.05; Values are SF36Vitality change per measurement occasion.
Random intercept only model, linear growth – linear growth model (random intercept random slope), quadratic growth – quadratic growth model, Spline 1 – Linear splines with knot at 2nd measurement occasion (6months), Spline 2 – Spline with knot at 2nd (1 year) measurement occasions.

The Linear growth model is selected as the best fit for the data, with better fit statistics. Observations represent the total number of entries in the dataset, comprising 729 participants, and five measurement occasions.

Appendix Table 11.7: Unconditional and conditional growth model unimputed data

Unconditional model				Fully conditional model		
Predictors	β (CI)	Std β (CI)	p	B (CI)	Std (β)	p
(Intercept)	42.54 (41.78 – 43.31)	-0.01 (-0.08 – 0.05)	<0.001	21.89 (18.90 – 24.88)	0.07 (0.00 – 0.14)	<0.001
visitx	0.17 (0.03 – 0.32)	0.03 (0.00 – 0.06)	0.021	-0.15 (-0.26 – -0.03)	-0.03 (-0.05 to -0.01)	0.014
Age				0.01 (-0.03 – 0.04)	0.01 (-0.03 – 0.05)	0.753
Sex				-1.28 (-2.26 – -0.30)	-0.11 (-0.19 to -0.03)	0.011
HAQ				-2.73 (-3.32 – -2.14)	-0.18 (-0.22 to -0.14)	<0.001
SF36BP				0.22 (0.18 – 0.26)	0.2 (0.17 – 0.24)	<0.001
SF36MH				0.39 (0.36 – 0.42)	0.37 (0.34 – 0.40)	<0.001
PGA				-0.04 (-0.05 – -0.02)	-0.08 (-0.19 to -0.03)	<0.001
Random Effects						
σ^2	48.87			32.86		
τ_{00}	81.53 $_{cenid}$			31.11 $_{cenid}$		
τ_{11}	0.76 $_{cenid.visitx}$			0.29 $_{cenid.visitx}$		
ρ_{01}	-0.03 $_{cenid}$			-0.28 $_{cenid}$		
ICC	0.64			0.48		
N	729 $_{cenid}$			729 $_{cenid}$		
Observations	2886			2817		
Marginal R^2 / Conditional R^2	0.001 / 0.643			0.514 / 0.746		

Intercept – Vitality values at baseline, β – Beta coefficient, std. - standardised, CI - Confidence Interval. σ2- Residual Variance, τ00 – Intercept variance, τ11 – Variance due to measurement occasion, ρ01– Covariance due to participant and measurement occasion, ICC – Intraclass correlation coefficient, N – Total number of participants, Observations – number of measurements included in analysis. Vitality - SF36 Vitality, Age - Age at onset, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Pain - SF36 Bodily Pain, Mental health - SF36Mental health

Appendix Table 11.8: Unconditional and fully conditional Growth models, ERAN participants with five complete Vitality measurements and in people with disease duration of six months or less.

	Disease duration less than 3 months			**5 complete SF36 Vitality measurements**		
Predictors	Unconditional model	Conditional model without interactions	Conditional model with interactions	Unconditional model	Conditional model without interactions	Conditional model with interactions
	Estimates (CI)	Estimates (CI)	Estimates (CI)	Estimates (CI)	Estimates (CI)	Estimates (CI)
(Intercept)	42.47 (41.69 to 43.26) *	20.73 (18.07 to 23.40) *	22.14 (18.33 to 25.95) *	44.61 (43.19 to 46.03) *	23.51 (18.93 to 28.08) *	20.89 (15.01 to 26.77) *
Measurement Occasion	0.25 (0.09 to 0.41) *	-0.15 (-0.26 to -0.03) *	-0.71 (-1.75 to 0.34) *	0.08 (-0.16 to 0.33)	-0.22 (-0.40 to-0.03) *	0.82 (-0.83 to 2.47)
Sex		-1.39 (-2.27 to -0.51) *	-1.26 (-2.39 to 0.13) *		-1.05 (-2.65 to 0.56)	-1.24 (-3.05 to 0.57)
HAQ		-2.78 (-3.31 to -2.24) *	-3.10 (-3.88 to -2.32) *		-2.99 (-3.91 to -2.08) *	-2.68 (-3.85 to -1.52) *
SF36 MH		0.40 (0.38 to 0.43) *	0.39 (0.35 to 0.45) *		0.38 (0.33 to 0.43) *	0.42 (0.35 to 0.49) *
SF36 BP		0.26 (0.22 to 0.30) *	0.24 (0.18 to 0.29) *		0.25 (0.19 to 0.31) *	0.27 (0.18 to 0.35) *
PGA		-0.03 (-0.04 to -0.01*	-0.02 (-0.04 to -0.002) *		-0.04 (-0.06 to -0.02) *	-0.04 (-0.07 to -0.01) *
Occ:sex			-0.04 (-0.29 to 0.20)			0.11 (-0.31 to 0.52)
Occ:HAQ			0.13 (-0.09 to 0.35)			-0.15 (-0.49 to 0.20)
Occ:sfnMH			0.004 (-0.01 to 0.02)			-0.02 (-0.04 to 0.001)
Occ:sfnBP			0.01 (-0.01 to 0.03)			-0.01 (-0.03 to 0.02)
Occ:PGA			-0.002 (-0.01 to 0.01)			0.0004 (-0.01 to 0.01)

***P ≤0.05 Conditional model – Model with covariates with and without an interaction between covariates and measurement occasions Intercept – Vitality values at baseline. SF36VT - SF36 Vitalit, SF36MH - SF36 Mental Health, SF36BP - SF36 Bodily Pain, HAQ - Health Assessment Questionnaire, PGA- Patient's Global Assessment of Disease Activity, Occ:sex – Sex and time interaction, Occ:SF36MH – Mental Health and time interaction, Occ:SF36BP – Bodily Pain and time interaction, PGA and time interaction, Occ:PGA- - PGA and time interaction**

Appendix Table 11.8 represents estimates from linear growth models investigating the progression of fatigue over the study duration, in people with 5 complete Vitality measurements, and in people with a disease duration of less than 6 months. Estimates from both models are similar to the results from the total ERAN population.

Appendix Table 11.9 below shows the random effects of the models described above. Random effect parameters produced comparable results to those presented in the main analysis.

Appendix Table 11.9: Random effects of models in people with a disease duration of less than six months and participants with five complete measurements.

Parameters	**Disease duration less than 6 months**			**5 complete SF36 Vitality measurements**		
	Unconditional model	Model without interactions	Model with interactions	Unconditional model	Model without interactions	Model with interactions
σ^2	73.59	42.84	42.85	43.84	28.51	28.26
τ_{00}	64.53	22.13	22.31	88.11	22.84	22.67
τ_{11}	1.07	0.20	0.20	1.36	0.57	0.64
ρ_{01}	-0.30	-0.51	-0.52	-0.07	-0.16	-0.17
ICC	0.46	0.31	0.31	0.69	0.48	0.49
N	631	631	631	208	208	208
Observations	3155	3155	3155	1040	1040	1040
Marginal R^2 / Conditional R^2	0.002 / 0.466	0.542 / 0.682	0.541 / 0.681	0.000 / 0.694	0.578 / 0.781	0.577 / 0.782

σ2- Residual Variance, τ00 – Intercept variance, τ11 – Variance due to measurement occasion, ρ01– Covariance due to participant and measurement occasion, ICC – Intraclass correlation coefficient, N – Total number of participants.

Appendix Table 11.10: Fit statistics of GBTM models for people with disease duration of less than six months and five compete measurements

Number of groups	Less than 6 months				5 Complete Vitality measurements			
	BIC	AIC	LL	Entropy	BIC	AIC	LL	Entropy
1	-9,676.24	-9,669.56	-9,666.56		-4,063.21	-4,058.21	-4,055.21	
2	-9,241.45	-9,225.88	-9218.88	0.80	-3,830.31	-3,818.63	-3,811.63	0.89
3	-9,121.31	-9,101.30	-9092.30	0.79	-3,784.56	-3,764.54	-3,752.54	0.85

BIC – Bayesian Information Criteria, AIC – Akaike Information Criteria, LL – Log Likelihood.

Group selection fit statistics are presented in Appendix Table 11.10 for participants with a disease duration of less than 6 months and 5 complete Vitality measurements. Values select 2 rather than a 3 - group membership, with lower BIC and higher entropy.

Appendix Table 11.11:Fatigue Group membership in people with disease duration of less than 6 months, n=631

Groups	**Parameters**	**Estimate**	**95% CI**	**Group membership** (%)
Fatigue	Intercept	35.70	34.45 -36.00*	50.75
	Linear	-0.04	-0.15 - 0.29	
No Fatigue	Intercept	49.85	48.70 - 50.64*	49.45
	Linear	2.33	0.95 - 2.57*	
	Quadratic	-0.49	-0.37 to -0.11*	

Intercept – SF36VT values at baseline, Linear/Quadratic - SF36VT growth rate per measurement occasion, * P ≤ 0.05, Group membership - % of total population assigned to each group

Appendix Table 11.11 and 11.12 show parameter estimates from models examining trajectory groups in people with a disease duration of less than 6 months, and in people with 5 complete Vitality measurement. Estimates show that the proportion of people in the fatigue group was slightly higher in people with 5 complete measurements compared to the total population and those with a disease duration of less than 6 months.

Appendix Table 11.12: Fatigue Group membership in people with complete measurements(n=208)

Groups	**Parameters**	**Estimate**	**95% CI**	**Group membership** (%)
Fatigue	Intercept	35.70	34.45 - 36.00*	43.23
	Linear	-0.43	-0.15 - 0.29	
No Fatigue	Intercept	50.08	48.70 - 50.64*	56.77
	Linear	4.26	0.95 - 2.57*	
	Quadratic	-1.20	-0.37 to -0.11*	

Intercept – SF36VT values at baseline, Linear/Quadratic - SF36VT growth rate per measurement occasion, * P ≤ 0.05 Group membership - % of total **population**

Appendix Table 11.13: Number of groups selected through alternative model specification

	loglik	AIC	SABIC	entropy	%class1	%class2	%class3
1 Group	-10527.44	21066.88	21075.38		100		
2 Groups	-10518.57	21055.14	21067.89	0.49	45.8161866	54.18381	
3 Groups	-10515.56	21055.13	21072.12	0.64	0.9602195	44.99314	54.04664

Appendix Table 11.13 above shows group selection using an alternative specification (Growth Mixture Modelling). The GMM model favours a 2-group membership, with a lower AIC, BIC, and log link. Although the 3-group model has a larger entropy, the 2 -group model is superior because in addition to other fit statistics mentioned earlier, the proportion of individuals in class 1 is less than 1% of the population, favouring an alternative grouping structure. Finally, estimated proportions in each of group of the 2-group model is similar to estimates obtained using group-based trajectory modelling.

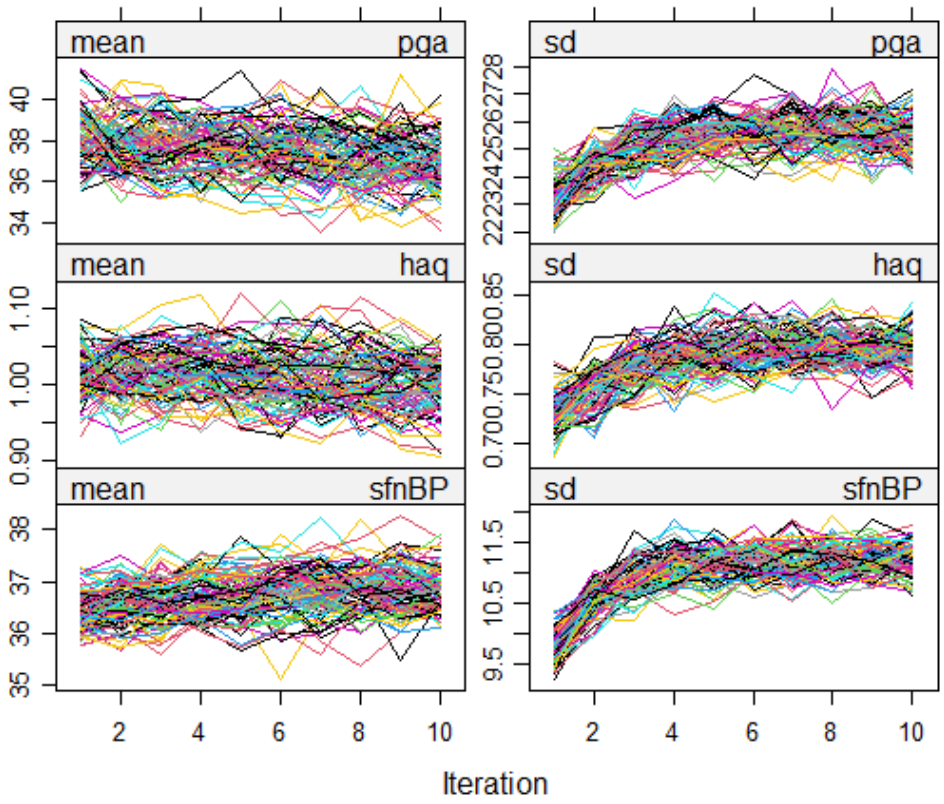


Figure 11.10: Multiple imputation convergence plots

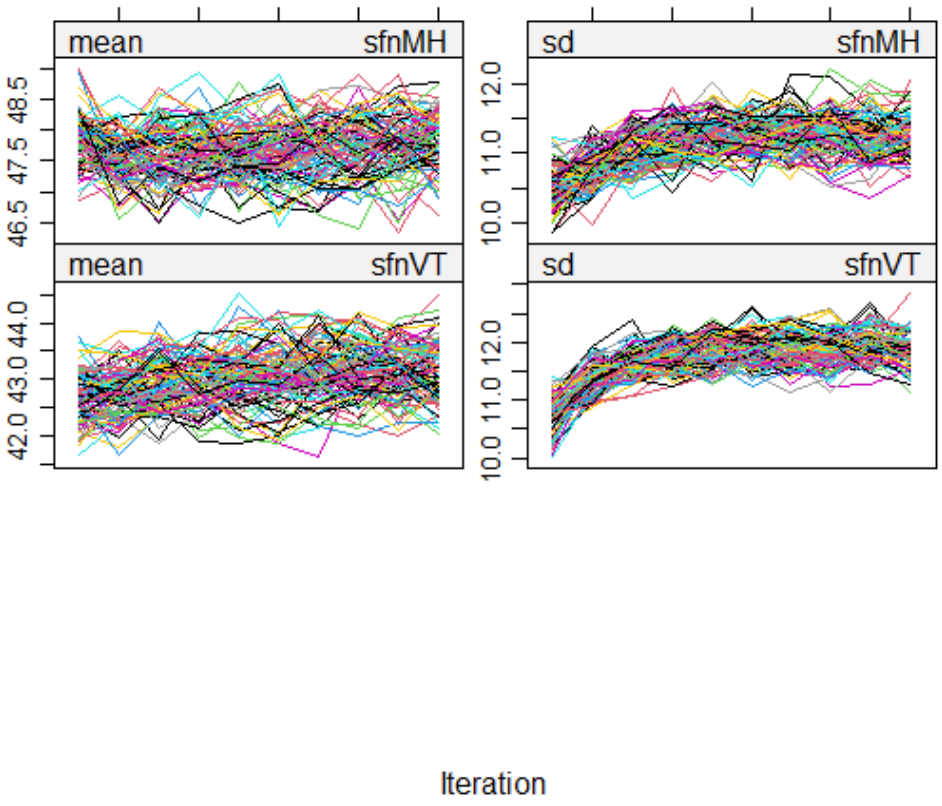
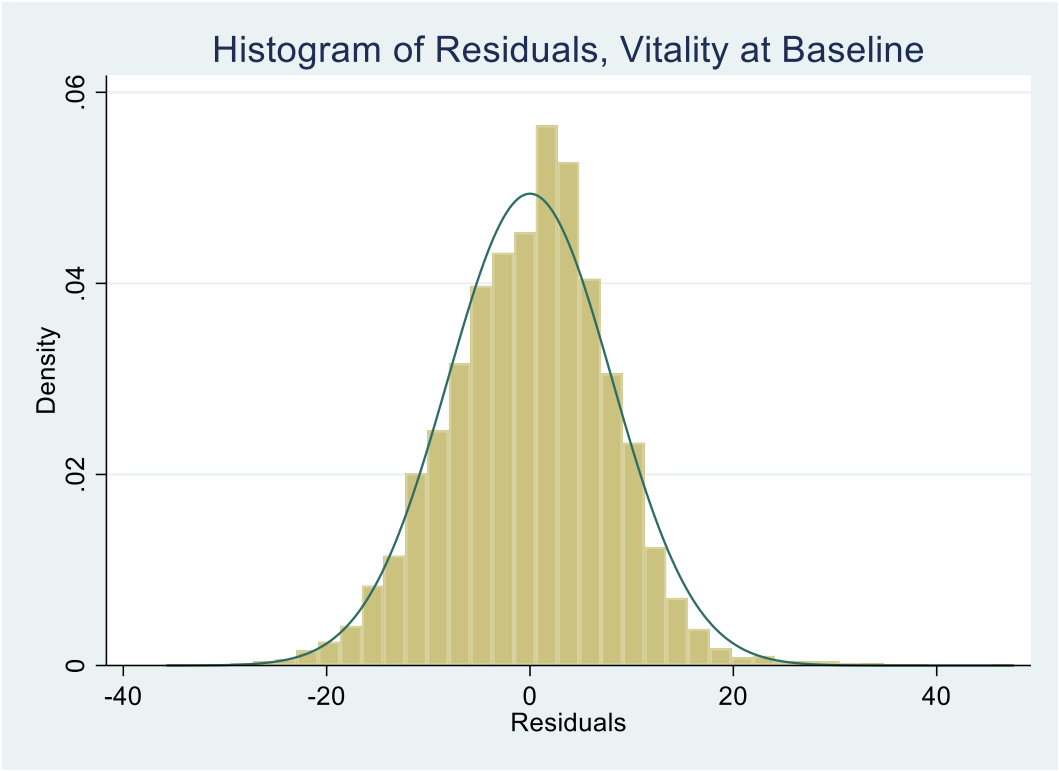
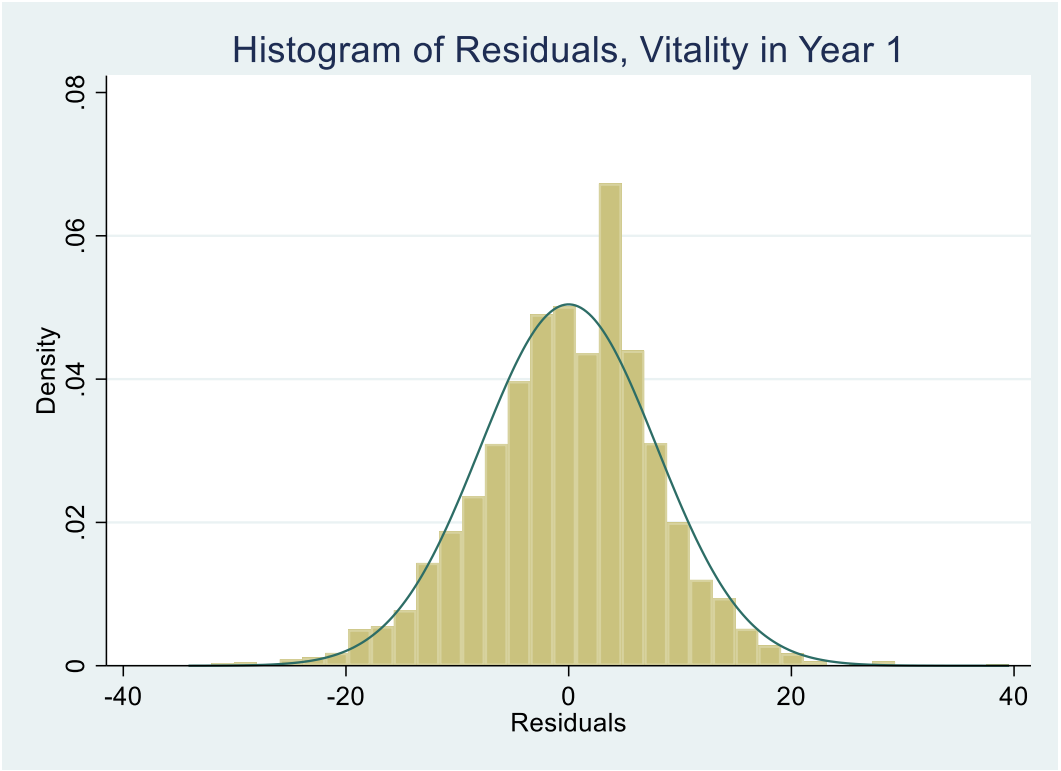


Figure 11.11: Multiple imputation convergence plots 2

11.6 KPIC study appendices



Appendix Figure 11.12:Histogram of residuals of the multivariable linear regression model of the characteristics associated with SF12Vitality at baseline.



Appendix Figure 11.13: Histogram of residuals of the multivariable linear regression model of the characteristics associated with SF12Vitality in year 1.

11.6.1 Weighted KPIC population analysis (Estimating the effects of non-response bias)

KPIC population was compared with the East Midlands population obtained from the Office of National Statistics (ONS), https://api.beta.ons.gov.uk/v1/datasets/create/filter-outputs/e8eb2bf8-cee3-45c3-8210-35b5066b1b7e. Census data were categorized by age and sex according to the structure of KPIC data to allow for comparison. Proportions of each category was compared using the equality of proportions test (Table 11.14). KPIC was weighted to the census population using iterative proportional fitting (procedure described elsewhere). Table 11.15 shows the distribution of the weighted KPIC population. Tables 11.16 – 11.18 show the results of linear and logistic regression analysis estimating the association between fatigue and covariates in the weighted and unweighted KPIC populations. Findings generally show very similar point estimates and confidence intervals, with, the model covariates and direction of association remaining the same in both datasets. This is because the distribution of covariates in the weighted and unweighted datasets are very similar (Table 11.15).

11.6.1.1 Weighted KPIC population analysis (Estimating the effects of non-response bias)

Table 11.14:Comparing East midlands population and the KPIC sample

Characteristic	Group	census	KPIC	Difference	*CI	pvalue
Age	≤49	0.24	0.15	0.09	0.08 to 0.09	<0.0001
	50 - 59	0.27	0.24	0.039	0.02 to 0.04	<0.0001
	60 - 69	0.22	0.33	-0.11	-0.12 to -0.10	<0.0001
	≥70	0.27	0.28	-0.01	-0.02 to 0.002	0.139
Sex	female	0.52	0.57	-0.05	-0.06 to -0.04	<0.0001
	male	0.48	0.43	0.05	0.04 to 0.06	<0.0001

***95% Confidence interval (CI)and P values obtained using a 2 sample equality of proportions test. Values are proportions**

Appendix table 11.15: Comparing Weighted population with KPIC population

		N - 9, 513		
Characteristics	**unit**	**n**	**Unweighted_sample**	**Weighted Descriptives**
Demographics				
Age mean(sd)	years	9,452	62.09 (10.58)	60.03 (11.49)
Age categories				
BMI median(IQR)	kg/m2	9,070	26.45 (23.76 - 29.70)	26.47 (23.80 - 29.76)
Sex	male	4,090	43.26%	48.33
	female	5,365	56.74	51.67
Fatigue				
Fatigue	yes	5,944	63.36*	64.09*
	No	3,437	36.64*	35.91*
SF12Vitality	0 -100	9,406	43.39(11.51)	43.41 (11.50)
Pain				
SF12 Bodily Pain	0-100	9,380	41.00 (10.37)	41.16 (10.34)
Knee Pain	yes	2,686	45.3	44.5
	No	3,244	54.7	55.5
Mental Health				
HADs- Anxiety	0 -21	9,418	6 (3 - 9)	6 (3 - 9)
HADs-Depression	0-21	9,422	3 (2-6)	3 (2-6)
Sleep				
Sleep Adequacy	0-100	9,370	57.50 (29.54)	56.97 (29.71)
Laboratory				
CRP	mg/l	213	2.02 (1.05 - 4.45)	1.97 (1.01 - 4.49)
Comorbidity	count (0-7)	9,513	1 (0 - 2)	1 (0 - 2)
Physical Activity		n	%	
GPPAQ	Inactive	86	20.77	23.47
	Mod. Inactive	94	22.71	22.65
	Mod. active	66	15.94	15.64
	Active	168	40.58	38.24

Table 11.16: Bivariate analysis SF12Vitality, KPIC population vs weighted population

Characteristics	unit	Total Population				Weighted Population		
		ß	CI	p value	n	β	CI	p value
Age	years	0.01	-0.02 to 0.03	0.661	9,350	0.01	-0.02 to 0.03	0.505
BMI	kg/m2	-0.52	-0.57 to -0.48	<0.001	8,976	-0.50	-0.56 to -0.46	<0.001
	Female	-1.95	-2.41 to -1.48	<0.001	9,355	-2.06	-2.56 to -1.57	<0.001
SF12 Bodily Pain	0 -100	0.6	0.58 to 0.62	<0.001	9,342	0.59	0.57 to 0.61	<0.001
Knee Pain	Yes	-6.12	-6.70 to -5.54	<0.001	5,862	-6.12	-6.74 to -5.50	<0.001
HADs- Anxiety	0 -21	-1.32	-1.37 to -1.27	<0.001	9,354	-1.32	-1.37 to -1.26	<0.001
HADs- Depression	0-21	-2.03	-2.08 to -1.99	<0.001	9,350	-2.00	-2.06 to -1.95	<0.001
Sleep	0-100	0.18	0 .17 to 0.19	<0.001	9,303	0.18	0.17 to 0.18	<0.001

ß – Regression coefficient, CI – 95% Confidence Interval

Appendix table 11.17: Multivariable analysis SF12 Vitality, KPIC population vs weighted population

	n = 8,777		**R^2 = 0.50**			**R^2 = 0.50**			
		Total Population				Weighted Population			
Characteristics	Unit	ß	CI	p value	Std ß	ß	CI	p value	Std ß
Bodily pain	0 – 100	0.22	0.20 to 0.24	<0.001	0.20	0.21	0.18 to 0.23	<0.001	1.8
Depression	0 - 21	-1.30	-1.41 to -1.25	<0.001	-0.42	-1.30	-1.36 to -1.23	<0.001	-0.42
Anxiety	0 - 21	-0.11	-0.15 to -0.03	0.003	-0.03	-0.12	-0.17 to -0.06	<0.001	-0.42
Sleep score	0 – 100	0.07	0.06 to 0.07	<0.001	0.17	0.064	0.06 to 0.07	<0.001	0.17
BMI	Kg/M^2	-0.12	-0.16 to -0.06	<0.001	-0.06	-0.11	-0.15 to -0.08	<0.001	-0.05
Comorbidities	0 – 7	-0.92	-1.10 to -0.64	<0.001	-0.08	-0.83	-1.00 to -0.65	<0.001	-0.07
Female	Yes	-1.05	-1.40 to -0.70	<0.001	-0.05	-1.20	-1.55 to -0.85	<0.001	-0.05
All $p \leq 0.05$, ß – Regression coefficient, Std ß = Standardized regression coefficient, CI – 95% Confidence Interval									

Appendix table 11.18 Bivariate analysis Fatigue, KPIC population vs weighted population

		Total Population				Weighted Population		
Characteristics	unit	OR	CI	P-value	n	OR	CI	P-value
Age	years	0.98	0.98 to 0.99	<0.001	9,323	0.98	0.98 to 0.99	<0.001
BMI	kg/m2	1.07	1.06 to 1.08*	<0.001	8,951	1.07	1.06 to 1.08	<0.001
Sex	Female	1.60	1.47 to 1.74	<0.001	9,328	1.64	1.50 to 1.79	<0.001
Pain*	0 -100	0.38	0.36 to 0.40	<0.001	9,283	0.39	0.37 to 0.41	<0.001
Anxiety	0 -21	1.30	1.29 to 1.33	<0.001	9,320	1.31	1.29 to 1.33	<0.001
Depression	0-21	1.44	1.41 to 1.47	<0.001	9,324	1.43	1.40 to 1.46	<0.001
Sleep*	0 -100	0.65	0.64 to 0.67	<0.001	9,350	0.66	0.64 to 0.67	<0.001
Comorbidity	count (0-7)	1.45	1.38 to 1.51	<0.001	9,381	1.40	1.34 to 1.473	<0.001

OR – Odds ratio, CI – 95% Confidence Interval, *ORs for every 10-unit increase.

Appendix Table 11.19Multivariable analysis Fatigue, KPIC population vs weighted population

n=8,763							
Characteristics	unit	Total Population			Weighted population		
		OR	CI	p value	OR	CI	p value
Age	years	0.99	0.98 to 0.99	<0.001	0.99	0.98 to 0.99	<0.001
BMI	kg/m2	1.02	1.01 to 1.03	0.001	1.02	1.01 to 1.04	0.001
Sex	Female	1.32	1.23 to 1.46	<0.001	1.38	1.23 to 1.55	<0.001
Pain*	0-100	0.65	0.60 to 0.70	<0.001	0.66	0.61 to 0.71	<0.001
Anxiety	0 -21	1.11	1.09 to 1.12	<0.001	1.11	1.09 to 1.13	<0.001
Depression	0-21	1.18	1.15 to 1.22	<0.001	1.18	1.14 to 1.21	<0.001
Sleep*	0-100	0.74	0.72 to 0.76	<0.001	0.74	0.72 to 0.76	<0.001
Comorbidity	0 - 7	1.21	1.14 to 1.32	<0.001	1.20	1.12 to1.28	<0.001
OR – Odds ratio, *ORs for every 10-unit increase.							

11.7 CAPRA study appendices

11.7.1 SPIRIT checklist

SPIRIT 2013 Checklist: Recommended items that are addressed within the CAP-RA study protocol.

Section/item	Item No	Description	Tick
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	✓
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	✓
	2b	All items from the World Health Organization Trial Registration Data Set	NA
Protocol version	3	Date and version identifier	✓
Funding	4	Sources and types of financial, material, and other support	✓
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	✓
	5b	Name and contact information for the trial sponsor	✓
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	✓
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	✓
Introduction			
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	✓
	6b	Explanation for choice of comparators	NA
Objectives	7	Specific objectives or hypotheses	✓
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	✓
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	✓
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	✓
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	NA
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	NA
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	NA

Section/item	Item No	Description	Tick
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	NA
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (e.g., change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	✓
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	✓
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	✓
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	✓
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	NA
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	NA
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	NA
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	NA
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	NA
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	✓
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	✓
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	✓
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	✓
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	✓
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	✓
Methods: Monitoring			

Section/item	Item No	Description	Tick
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	✓
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	✓
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	✓
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	NA
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	✓
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	✓
Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	✓
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	✓
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	✓
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	✓
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	✓
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	NA
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	✓
	31b	Authorship eligibility guidelines and any intended use of professional writers	NA
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	✓
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	✓
Biological specimens	**33**	**Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable**	✓

Checklist of SPIRIT guideline items, taken from SPIRIT template for clinical trials. Ticked items are either reported within this manuscript or detailed in the study protocol. NA = not applicable for the CAP-RA study.

11.7.2 CAP-RA Steering Committee TOR

CAP-RA Study Steering Committee Terms of Reference

The terms of reference define the objectives, operational scope, outlines the organization and working methods of the Central Aspects of Pain in Rheumatoid Arthritis (CAP-RA) study steering committee.

Remit

The CAP-RA steering committee provides strategic advice to the chief investigator (CI) /principal investigator (PI) and members of the research team on key areas of the study, including but not limited to:-

1. Monitor and supervise of all aspects of the study to ensure that the interim and overall objectives of the study are met.
2. Meet frequently, at least quarterly throughout the duration of the project to discuss CAP-RA related issues
3. Monitor study progress against pre-agreed milestones (such as recruitment and follow up rates), adherence to agreed protocol.
4. Advice on proposed changes to study plans in light of new evidence or unanticipated developments e.g major and minor ethics amendments, pandemics etc.
5. Provide advice to funders, investigators, study sponsor and other stakeholders on the progression of the project through reports or minutes of meetings, including the end-of-project reports.
6. Oversee and support any requests for additional funding or time extensions.
7. Oversee dissemination of research findings to research participants and general public including submission to conferences and peer-reviewed journals.
8. Oversee end of project report writing

Membership

9. The CAP-RA steering committee will comprise of members of the core research team including patient members.
10. The steering committee will be chaired by the CI/PI or a delegated member of the committee, who will have the responsibility for the timetable, agenda and conduct of the steering committee meetings. Other members of the committee will support the chair to carry out his/her responsibilities.
11. A deputy chair will be appointed and will be responsible for chairing meetings in the absence of the chair.
12. A secretary will be appointed by the team to oversee team administration e.g scheduling meetings, minute writing and circulation.
13. Membership of the steering committee will be nominated and agreed by existing members of the steering committee.
14. All members will be expected to commit for the duration of the study.
15. Should a member need to step down from the steering committee, an appropriate replacement will be invited to join following guidelines listed in (13) above.
16. Membership of the CAP-RA steering committee will be kept lean to facilitate effective and focused discussions.

Attendees

17. The study sponsor/funder, representatives of the host institution, and regulatory agencies relevant to the study may attend meetings as observers.
18. Other attendees may be invited as guests where their expertise is required but this does not confer automatic membership to the committee.

Accountability

19. The CAP-RA steering committee acts in an advisory capacity and is accountable to the CI/PI and study sponsor, who make all decisions regarding the delivery of the study.

Meetings

20. The steering committee is expected to meet every quarter for the duration of the study
21. Documents for the meeting will be sent to members in a timely fashion, at least one week prior to the meeting date.
22. At least half of the members including the chair or designated deputy chair must participate in a meeting either in person or via teleconference, or other virtual meeting avenues.

Minutes and Reporting

23. Secretariat and administration of the CAP-RA steering group will be provided by the University of Nottingham
24. The minutes of the meeting will be circulated to all CAP-RA steering committee members and may be available to other stakeholders as appropriate.

Conflicts of interest

25. Members of the CAP-RA steering committee will be required to inform the CI/PI of any potential or actual conflicts of interest arising in relation to the study. Interests should be declared as soon as they arise, at each setting giving rise to the conflict.
26. All declared interests will be reviewed by the CI/PI, who will then decide whether, given the circumstances of the interest, the individual should:
    a. Contribute/take part in discussions as appropriate
    b. Remain privy to information, but not contribute to activities or take part in any discussion specific to the conflict of interest, except to answer any relevant questions that might reasonably be put to them.
    c. Not be privy to information specific to the conflict of interest
27. If the CI/PI declares a conflict, the deputy chair will review the circumstances and make the decision. In chair and deputy chair declare the same conflict, a member of the committee will make the decision.

Expenses

28. Expenses/fees for travel and subsistence will be paid in line with University of Nottingham policies, as relevant

Confidentiality

29. Members of the CAP-RA steering group will agree to maintain confidentiality regarding items discussed at the meetings

Review

30. Terms of reference of Reference will be reviewed as required.

Tenure

31. The Steering Committee will disband when the study has been reported closed

List of members 1st April 2020:

Professor David Walsh

Professor Eamonn Ferguson

Dr Daniel McWilliams

Dr Onosi Ifesemen

Mrs Dorothy Plats

Mrs Susan Ledbury

Mrs Deborah Wilson

11.7.3 Minutes from PPI Focus Group - Kings Mill Hospital 19 February 2020

Central Aspects of Pain in Rheumatoid Arthritis (CAP-RA) Questionnaire

Present:

Phil Buckley, Debbie Wilson, Julie Jones-Diette (minutes), Onosi Ifesemen (Research team representative), Dorothy Platts (Steering group PPI rep), Suze Ledbury (Steering group PPi rep), Linda Waterfield, Maureen Voce, Donner Butler, Rosie Dyce, Nyreen Wallett, Paula Cooke.

JJD - Introduction and explanation of event and study to the coordination team (DW, PB) and steering group representatives (DP, SL) prior to arrival of focus group participants.

OI – Explanation of research study background, aims and objectives.

JJD – Welcome to focus group participants and explanation of the plan for the day.

OI – Welcome and introduction, aim of focus group and what will be the aim of the day. Distribution of survey questions and time allowed to review the survey and answer the questions from own experience. Discussion followed facilitated by OI.

Discussion: (*in italics*)

" Does the question convey the right message to the participant?"

1. **Temperature: `Cold or heat (e.g. bath water) on my joint was painful'**

This item examines the presence of hypersensitivity to heat or cold on painful joints. i.e application of a cold or hot stimulus that would normally not cause pain when in physical contact with unaffected joints.

Cold can be bad but warm can be soothing

I don't have a bath I only have showers – this was discussed the use of bath water temp as an example.

Temperature effects on the pain experience was the focus of the discussion and how best to apply heat of cold to the painful area.

2. **Fatigue: 'I generally felt tired'**

This item investigates the enduring, subjective sensation of generalized tiredness or exhaustion, which may or may not be related to joint pain or RA pain.

Fatigue is one of the most awful feelings I can hardly stay awake sometimes. As soon as I stop I feel like falling asleep

I don't stop that way I don't feel fatigued I get up early and I keep going

I do too I work and I'm very busy but as soon as I stop the fatigue is so overwhelming

I just don't stop I get up at 5am and I keep going until 8pm

Fatigue is very difficult to manage

3. **Cognitive impact: 'My joint pain stopped me concentrating on what I was doing'**

RA pain that interferes with, interrupts, distracts, or prevents someone from concentrating on something they were previously focused on.

I use to be thinking about my pain all the time wishing it would go away

Some of the group commented that they don't feel pain so this question was not relevant

Some discussed how difficulty it can be to just do normal tasks like gardening. Comment was made describing the payoff I go ahead and do what I need to but I know that I will be in pain and will need to rest all the next day.

At this point the group were brought back to focus on the meaning of the question by JJD rather than answering the question from their own experiences. The group were asked to comment on what they think the question is asking, what is the intended meaning and when they read the question what did it make them think?** see below

It's the little things that are frustrating the things that are difficult like showering. I cannot dry myself very well which means it is very difficult getting dressed you need to find alternatives to doing very simple tasks like putting underwear (bra) on.

When I read the question I thought if something hurts to do then its time to stop you have to listen to your body.

I think I'm my own worst enemy because I just won't stop

4. **Catastrophizing: 'I kept thinking about how much my joint hurt'**

This item investigates the focused attention on the painful joints, on its possible causes and consequences, as opposed to its solutions - ruminative thoughts.

This item seeks to identify the frequency of ruminative thoughts about having joint pain, and the RA pain experience.

I think this is asking how much is your joint hurting and are you fixating on it. It's important to try not to focus on the pain not to fixate on it otherwise you can't stop thinking about it.

5. **Anxiety: 'In general, I got sudden feelings of panic'**

This item seeks to identify whether the respondent has had feelings of panic, nervousness, or unease, which might or might not be related to RA joint pain.

This made me think about worrying about how will I get my clothes on and who will help me if I need help in my own bathroom when I won't let anyone in. This makes me panic.

Made me think about what if the drugs stop working or if they take them away if they withdraw my drugs that makes me panic.

This question is about asking if you are coping with tasks, pain medication, managing your pain.

I don't panic, the only time I panic is when I had shoulder surgery and I worried about how to get my underwear on. – A discussion followed in detail about methods to help with underwear detail not captured as not relevant.

The group agreed that generally they don't experience feelings of anxiety or panic.

6. **Sleep: 'My joint pain affected my sleep'**

This item investigates the frequency of RA pain disrupting sleep (nocturnal or daytime), including poor quality of sleep, insufficient or too much inefficient sleep, or interrupting sleep cycle.

I dread going to bed. I can't get comfortable, and I can't move easily to get into a good position its so uncomfortable and I can't sleep. I sleep in a recliner (chair) now because it's the only way I can get some rest.

I don't think that's good studies show you should try to maintain sleeping in your bed and not sleeping in a chair, you need to keep the routine.

I prefer to sleep in my chair it's been 7 or 8 weeks since I slept in my bed.

7. **Depression: 'I generally still enjoyed the things I used to enjoy'**

This item examines the participants experience of a loss of interest or pleasure. Enjoyment is the feeling of pleasure, satisfaction, or happiness that you have when you do or experience something that you like.

Yes

It lives with me I don't live with it

The pain in early diagnosis was just awful I felt like dying

(in my early diagnosis stage) I don't feel like dying

Well I did it was awful the earlier experiences and the pain are worse than when you have lived with it for years.

Yes it gets better when you first have it it can be very bad but it gets better and you learn to cope with it better.

8. **Pain distribution: 'This final question is about pain you may have had in any part of your body. Please shade in the diagram below, to indicate where you have suffered any pain for most days in the last 4 WEEKS. By pain we also mean aching and/or discomfort. Please do not include pain due to feverish illness such as flu'**

   This item investigates location of pain, ache, or discomfort across anatomical sites.

It was agreed the group understood that they were to shade the areas they felt pain not for other reasons.

OI asked for additional feedback on the collection of information weekly via text messaging. A handout was circulated to the group containing an example of the type of message a patient may receive.

A discussion followed,

Yes it is similar to the assessment by the specialist nurse asking about pain on a scale so easy to respond on a scale of 1 to 10.

Happy to respond.

OI - When should we send the question?

Responses would be different depending on the time of day.

Is the timing of the response important?

People may not respond at the time they get the message as they will be busy.

I don't have my phone with me at work.

I'm better first thing otherwise I will forget if I leave it

As the question is asking how much pain do you experience over the last week does it matter when we respond? It shouldn't matter when we respond as it's a average over the week. The timing of when to send the text doesn't matter.

Feedback from PPI on focus group event:

The group felt the experience was very good and thanked the organisers for creating the opportunity to get together.

The group asked one PPI participant who has recently been diagnosed if they had any feedback as they are new to the disease and PPI. This participant said they felt anxious about attending as they didn't really know what to expect but had found it to be a positive experience.

All participants would like to join the advisory group database, and all took a pack home.

The steering group PPI members felt the activity was very valuable and interesting.

Lessons learned:

The group received instructions verbally by phone call ahead of the meeting through discussion with JJD then again when they received the PIS and invitation letter via email ahead of the meeting and again at the start of the group activity as an introduction and again throughout the activity. This appeared to have little impact. The group was selected as a convenience sample at short notice, I think the lesson is that the value of this type of activity can be increased with additional planning, recruitment training and pre-work.

I noted many of the group had a lot of difficulty writing and as such completing the short questionnaire was a challenge.

** The group were not trained in how to participate in a focus group situation. As such they did not listen to the instructions and used the opportunity to talk about themselves and their own experiences.

Some members of the group were not suitable for inclusion into a focus group environment they were too vocal and used the opportunity as a platform to push their own views on others and often shut others down from expressing their views. They did not allow for full discussion.

11.7.4 CAP RA study materials

11.7.4.1 CAP-RA study poster

Dedicated to Outstanding care

Sherwood Forest Hospitals NHS Foundation Trust

University of Nottingham UK | CHINA | MALAYSIA

IRAS Project ID: 269143



CAP-RA STUDY

- **Do you have Rheumatoid Arthritis?**
- **Have you suffered from pain or fatigue?**
- **Are you interested in helping with research?**

The CAP-RA study is looking at ways to identify people for whom future treatments might prevent chronic pain, and persistent fatigue in Rheumatoid Arthritis.

For more information

Study webpage, tiny url or QR code

Study email and telephone number

For more information Study webpage, tiny url or QR code Study email and telephone	For more information Study webpage, tiny url or QR code Study email and telephone	For more information Study webpage, tiny url or QR code Study email and telephone	For more information Study webpage, tiny url or QR code Study email and telephone	For more information Study webpage, tiny url or QR code Study email and telephone	For more information Study webpage, tiny url or QR code Study email and telephone	For more information Study webpage, tiny url or QR code Study email and telephone	For more information Study webpage, tiny url or QR code Study email and telephone number

CAP-RA Poster final version 1.1 24/02/2020

11.7.4.2 CAP-RA Ethics Application Response

North of Scotland Research Ethics Committee (1)
Summerfield House
2 Eday Road
Aberdeen
AB15 6RE

Telephone: 01224 558458
Facsimile: 01224 558609
Email: nosres@nhs.net

NHS Grampian

Please note: This is an acknowledgement letter from the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

30 March 2020

Professor David A Walsh
Professor of Rheumatology
University of Nottingham
Academic Rheumatology, Clinical Sciences Building
City Hospital
NOTTINGHAM
NG5 1PB

Dear Professor Walsh

Study title:	**Improving Pain Outcomes in Rheumatoid Arthritis; Detecting the Contribution of Central Pain Mechanisms.**
REC reference:	**20/NS/0036**
Protocol number:	**20001**
IRAS project ID:	**269143**

Dr Ifesemen kindly e-submitted revised documents on 30 March 2020. I can confirm the REC has received the documents listed below and that these comply with the approval conditions detailed in our letter dated 27 March 2020.

Documents received

The documents received were as follows:

Document	Version	Date
IRAS Checklist XML [Checklist 30/03/2020]		30 March 2020
Letters of invitation to participant [CAP-RA Participant Invitation Letter]	final version 1.2	30 March 2020
Participant consent form [CAP-RA Consent Form]	final version 1.2	30 March 2020

Approved documents

The final list of approved documentation for the study is therefore as follows:

Document	Version	Date
Contract/Study Agreement template [UoN NCA Interventional Studies data sharing]	4.1	24 February 2020
Copies of advertisement materials for research participants [CAP-RA Poster]	final version 1.1	24 February 2020
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Evidence of Insurance]		30 July 2019
IRAS Application Form	269143/141 1513/37/77 8	17 February 2020
IRAS Checklist XML [Checklist 30/03/2020]		30 March 2020
Letter from funder [22462 - Walsh Notification]		23 December 2019
Letter from sponsor		03 February 2020
Letters of invitation to participant [CAP-RA Reliability Invitation Letter]	final version 1.1	24 February 2020
Letters of invitation to participant [CAP-RA Follow-up Invitation Letter 1]	final version 1.1	24 February 2020
Letters of invitation to participant [CAP-RA Follow-up Invitation Letter 2]	final version 1.1	24 February 2020
Letters of invitation to participant [CAP-RA Participant Invitation Letter]	final version 1.2	30 March 2020
Other [CAP-RA DAS28 CRF]	final version 1.0	13 February 2020
Other [CAP-RA Delegation Log]	final version 1.0	24 February 2020
Other [Pfizer General Research Grant Agreement]		20 September 2019
Other [No Opinion Letter]		11 March 2020
Participant consent form [CAP-RA Consent Form]	final version 1.2	30 March 2020
Participant information sheet (PIS) [CAP-RA Participant Information Sheet]	final version 1.1	24 February 2020
Referee's report or other scientific critique report [22462 Walsh Notification (Funder Letter)]		23 December 2019
Research protocol or project proposal [CAP-RA Protocol]	final version 1.1	24 February 2020
Summary CV for Chief Investigator (CI) [David Walsh]		24 February 2020
Summary CV for student [Onosi Ifesemen]		24 February 2020
Summary CV for supervisor (student research) [David Walsh]		24 February 2020
Summary CV for supervisor (student research) [Daniel McWilliams]		24 February 2020

Validated questionnaire [CAP-RA Baseline Questionnaire Booklet]	final version 1.0	13 February 2020
Validated questionnaire [CAP-RA Follow-up Questionnaire Booklet]	final version 1.0	13 February 2020
Validated questionnaire [CAP-RA Reliability Questionnaire]	final version 1.0	13 February 2020

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

20/NS/0036	**Please quote this number on all correspondence**

Yours sincerely



Ms Sarah Lorick
Assistant Ethics Co-ordinator

Copy to: Dr Onosi Ifesemen, Student Researcher
Alison Steel, Sherwood Forest Hospitals NHS Foundation Trust,
Lead Nation - England: approvals@hra.nhs.uk

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

Professor David A Walsh
Professor of Rheumatology
University of Nottingham
Academic Rheumatology, Clinical Sciences Building
City Hospital
Nottingham
NG5 1PB

Email: approvals@hra.nhs.uk
HCRW.approvals@wales.nhs.uk

14 April 2020

Dear Professor Walsh

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Improving Pain Outcomes in Rheumatoid Arthritis; Detecting the Contribution of Central Pain Mechanisms.**
IRAS project ID:	**269143**
Protocol number:	**20001**
REC reference:	**20/NS/0036**
Sponsor	**University of Nottingham**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) have been sent to the coordinating centre of each participating nation. The relevant national coordinating function/s will contact you as appropriate.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?

The standard conditions document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **269143**. Please quote this on all correspondence.

Yours sincerely,
Lauren Allen

Email: approvals@hra.nhs.uk

Copy to: *Dr Angela Shone*

11.7.4.3 CAP-RA Invitation Letter 1

Dedicated to Outstanding care

NHS Sherwood Forest Hospitals NHS Foundation Trust

University of Nottingham UK | CHINA | MALAYSIA

Central Aspects of Pain in Rheumatoid Arthritis (CAP-RA)

{INSERT PARTICIPANT ADDRESS}

Insert {Date}

Invitation to participate in a research study

Dear {INSERT PARTICIPAT NAME},

Pain and fatigue are common and sometimes severe in people with rheumatoid arthritis. Many people report pain and fatigue despite adequate treatment of their arthritis. We are therefore looking at some of the causes of pain and fatigue in people with Rheumatoid Arthritis.

In our other research we found that joint swelling and inflammation are not the only causes of pain and fatigue. The central nervous system (brain and spinal cord) can stop pain getting better in a process called central sensitization. Measuring central sensitization can be difficult and expensive.

We will like to invite you to take part in a study to develop a simple questionnaire to measure central sensitisation in people with Rheumatoid Arthritis. You are receiving this letter because you are under the care of the Rheumatology team at Sherwood Forest Hospitals NHS Foundation Trust, and may meet the eligibility criteria of the study. You may be contacted by telephone by a member of your clinical care team in about 1 week to follow up this letter. If you would not like to be contacted, please call this number [insert number of clinical team at SFH] or email [insert email of clinical care team].

Taking part in this study would involve two visit to Kings Mill Hospital, about 3 months apart. The visit will involve clinical assessments, blood tests, ultrasounds of your hands and feet, and you will be asked to complee some questionnaires. Through this study we hope to find better ways to measure central sensitisation and pain, to help us to develop treatments that better relieve rheumatoid arthritis pain and fatigue.

If you would like to take part in this study or have any questions regarding it, please contact the research team at [insert study phone number] or [insert study email address].

Please see the enclosed Participant Information Sheet for further information about our research project.

Thank you.

Yours sincerely,
Professor David Walsh
Consultant Rheumatologist, Sherwood Forest Foundation NHS Trust.
Chief Investigator

CAP-RA participant invitation letter final version 1.0 13/02/2020

11.7.4.4 CAP-RA Participant Information Sheet

Study Title: Central Aspects of Pain in Rheumatoid Arthritis (CAP-RA)
IRAS Project ID: 269143
Principal (Chief) Investigator and Study Doctor: Professor David Andrew Walsh, Consultant Rheumatologist Sherwood Forest Hospitals NHS Foundation Trust and Professor of Rheumatology, University of Nottingham.
Study Sponsor: University of Nottingham

Participant Information Sheet

You are being invited to take part in a research study by Sherwood Forest Hospitals NHS Foundation Trust and the University of Nottingham. The study is investigating pain in people with rheumatoid arthritis. Before you decide, it is important to understand why we are doing the study, how your information will be used, what the study will involve and the potential benefits, risks, and possible discomforts.

If you are willing to take part, you will be told about the research study and asked to initial and sign a consent form and will be given a copy to keep. If you give consent, the study may start immediately.

Why is this study being done?

Pain and fatigue are serious problems, even in people whose rheumatoid arthritis is well treated. In our other research, we have found that joint swelling and inflammation are not the only causes of pain. The central nervous system (brain and spine) does not merely sense pain; it also contributes to pain severity and can stop pain getting better. When the central nervous system makes pain much worse, we call this central sensitisation. Fatigue can be a part of this. The central nervous system also can sometimes block pain and make it feel better. Pain can affect how we think, feel and concentrate. Through this study, we hope to find better ways to measure central sensitisation and pain, to help us to develop better treatments. Parts of this study contribute to a PhD project.

Who can take part?

We are interested in everyone who has rheumatoid arthritis, even if pain and fatigue are not big problems for you. We hope that about 250 people will take part in this research study.

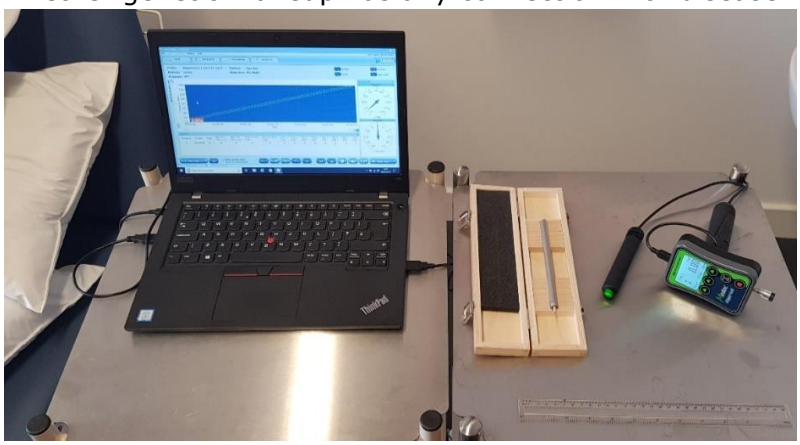
What are we asking you to do?

If you are interested in participating in this study, we would like you to come to King's Mill Hospital for clinical assessment, to complete questionnaire booklets, have ultrasound scans of your hands and feet, and give blood samples. These will be arranged at dates and times convenient for you. We may send you a short questionnaire a week after your first assessment. 3 - 4 months after the first visit, we would like to invite you to another visit. You will be asked to make 2 visits to King's Mill Hospital in total.

At each visit, you will see a member of our research team who has been specially trained for this study. Each study visit should be finished in 2-3 hours and will take place in a private clinic room.

During the study visit you will be asked to do the following things: -

Answer questions to check if you are suitable for this study, talk about the study and give your consent to participate
Discuss your medical records and information that could be relevant to arthritis and fill in a booklet of questions about your pain, health, sleep, medication and arthritis. This will include our new CAP-RA questionnaire. These are called "Study Data").
Have your joints examined for arthritis?
Do a 30 second memory test.
Be clinically assessed to see how you feel pain. Any pain that you feel during this assessment will be brief and should not be severe.
A researcher will use a softheaded probe on your forearm, knee, and leg to measure what is the lowest pressure that you feel as pain (pain detection threshold).
A researcher will assess your pain sensitivity by applying to your knee the tip of a specially manufactured pen.
A researcher will again assess your pain detection threshold with the soft-headed probe applied to your leg while your other arm is compressed by a blood pressure cuff. You will be asked to squeeze on a small soft ball (stress ball).
Have an ultrasound scan. A small handheld probe will be placed on and moved over the skin of your hand and foot. A lubricating gel is put on your skin to allow the probe to move smoothly. You shouldn't feel anything other than the touch of the probe on your skin, and the gel (which is often cold).
Provide a blood sample (which could be up to 30 ml or 6 tubes) which we will use to help measure inflammation and other molecules relevant to arthritis. Your blood might be used for genetic tests in order to determine whether genetic makeup has any connection with disease.



Pain assessment equipment
So that we can measure how much your answers might change, some people will be sent a second copy of the new CAP-RA questionnaire through the post a week after the study visit and asked to return the completed questionnaire in a pre-paid envelope.

We would also like to know how your fatigue and pain levels change from week to week. If you agree, you will receive 2 text messages each week for 12 weeks. The text messages will ask about your pain or fatigue levels during the past week on a scale of 0 – 10. You will be asked to reply to the text with a number. Text messages will be free of charge.

10 days before the date of your 2nd visit, you will receive a phone, text, email, or letter from the research team to confirm your appointment. You may receive another message a day before the study to remind you of the appointment. If due to some unforeseen circumstances you are unable to attend the second visit, you may be sent a questionnaire by post, to return in a prepaid envelope. You may contact the research team at any time during the study if you wish to reschedule your appointment.

Researchers at the University of Nottingham might wish to contact you with further information or to invite you to contribute further to their research. Receiving further information will not commit you to joining any studies.

What are the benefits of my taking part?
It is not intended for you to have direct personal gain from your participation. You will not be paid for taking part in this study. Travel expenses will be offered for any visits incurred as a result of participation (on provision of receipts and up to a maximum of £20).
Taking part in this study means that you may possibly help other arthritis sufferers in the future. Being able to measure the changes that occur in people with painful arthritic joints could be used to develop new ways of treating arthritis.
What are the risks if I take part?
We are not anticipating any significant risks to participants in this study. We appreciate that you are giving up your valuable time and will try to minimise the inconvenience. You will experience some mild and brief pain during the pain assessment. Everyone experiences pain sensations differently so if you ask us to stop the procedure, for any reason, we will immediately stop.

What if there is a problem?
If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. The researchers' contact details are given at the end of this information sheet. If you remain unhappy and wish to complain formally, you can do this by contacting the King's Mill Hospital patient advice and liaison service(PALS) at 01623 672222, email : PET@sfh-tr.nhs.uk.

In the unlikely event that something does go wrong, and you are harmed during the research and this is due to someone's negligence then you may have grounds for a legal action for compensation against the University of Nottingham but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Taking part is entirely voluntary.
It is up to you to decide whether or not to take part. If you do decide to take part, you will be given this information sheet and the consent form to sign and keep. If you decide to take part you are still free to withdraw at any time without giving any reason, and without your legal rights being affected. If you choose not to take part, now or in the future, you will not

be disadvantaged in any way, including your medical treatment and the care you are entitled to receive. You will be told if any new findings occur during the study that may affect your willingness to take part.
If you withdraw, we will no longer collect any information about you or from you but we will keep the information about you that we have already obtained as we are not allowed to tamper with study records and this information may have already been used in some analyses and may still be used in the final study analyses. To safeguard your rights, we will use the minimum personally identifiable information possible.

What will happen to any samples I give?
We would also like to seek your consent so that any remaining samples may be stored and used in possible future research – this is optional (please indicate on the consent form if you agree to this). The samples will be stored with a code unique to you and securely at the University of Nottingham under the University's Human Tissue Research Licence (no 12265). Some of these future studies may be carried out by researchers other than current team, who ran the first study, including researchers working for commercial companies. Any samples or data used will be anonymised, and you will not be identified in anyway. If you do not agree to this, any remaining samples will be disposed of in accordance with the Human Tissue Authority's codes of practice.

Please note that the link between you and the coded samples you have donated, and the associated medical data will be kept as long as your samples are retained. The link to your study data will be kept for at least 7 years after the end of the study. After this time the study data will be fully anonymised, as the link will be broken and so your Study Data can no longer be traced back to you.

Will taking part be confidential?
We will follow ethical and legal practice and all information about you will be handled in confidence.
If you join the study, we will use information collected from you and your medical records during the course of the research. This information will be kept **strictly confidential**, stored in a secure and locked office, and on a password protected database at the University of Nottingham. Under UK Data Protection laws the University is the Data Controller (legally responsible for the data security), and the Chief Investigator of this study (named above) is the Data Custodian (manages access to the data). This means we are responsible for looking after your information and using it properly. Your rights to access, change or move your information are limited as we need to manage your information in specific ways to comply with certain laws and for the research to be reliable and accurate. To safeguard your rights, we will use the minimum personally – identifiable information possible.

You can find out more about how we use your information and to read our privacy notice at:

https://www.nottingham.ac.uk/utilities/privacy.aspx.

The data collected for the study will be looked at and stored by authorised persons from the University of Nottingham who are organising the research. They may also be looked at by authorised people from regulatory organisations to check that the study is being carried out correctly. All will

have a duty of confidentiality to you as a research participant and we will do our best to meet this duty.

Where possible information about you which leaves the Hospital site will have your name and address removed and a unique code will be used so that you cannot be recognised from it, however sometimes we need to ensure that we can recognise you to link the research data with your medical records so in these instances we will need to know your name and date of birth. Your contact details will be held by the University for the duration of the study in order to arrange the study visits and send out questionnaires. If you agree to receive the text messages we will share your telephone number with the text messaging service, who will delete it at the end of the study.

Your contact information will be kept by the University of Nottingham for 15 years after the end of the study so that we are able to contact you about the findings of the study and possible follow-up studies (unless you advise us that you do not wish to be contacted). This information will be kept separately from the research data collected and only those who need to will have access to it. All other data (research data) will be kept securely for 7 years. After this time your data will be anonymised and retained. During this time all precautions will be taken by all those involved to maintain your confidentiality, only members of the research team given permission by the data custodian will have access to your personal data.

In accordance with the University of Nottingham's, the Government's, and our funders' policies we may share our research data with researchers in other Universities and organisations, including those in other countries, for research in health and social care. Sharing research data is important to allow peer scrutiny, re-use (and therefore avoiding duplication of research) and to understand the bigger picture in particular areas of research. Data sharing in this way is usually anonymised (so that you could not be identified) but if we need to share identifiable information, we will seek your consent for this and ensure it is secure. You will be made aware then if the data is to be shared with countries whose data protection laws differ to those of the UK and how we will protect your confidentiality.

Although what you say to us is confidential, should you disclose anything to us which we feel puts you or anyone else at any risk, we may feel it necessary to report this to the appropriate persons.

What will happen to the results of this study?
The results of the study may be published in medical literature and made more widely available, for example through our website (http://www.nottingham.ac.uk/paincentre), but your identity will not be revealed.

Who is organising and funding this study?
This study is organised and carried out by staff from the University of Nottingham and Sherwood Forest Hospitals NHS Foundation Trust. It is funded by Pfizer Ltd, Versus Arthritis and the National Institute for Health Research through grants to the University of Nottingham. The funders have no influence over the conduct of this study.

Who has reviewed this study?
All human research in the University of Nottingham is reviewed by independent groups of people, called a Research Ethics Committee to protect

your interest. All materials associated with this study have been reviewed and approved by the NHS Research Ethics Committee as well as the Health Research Authority.

Who can I contact if I have any questions about this study?
If you have any further questions about this study, or wish to contact the Study Sponsor, please discuss this in the first instance with:
Professor David Walsh, Principal (Chief) Investigator and Study Doctor**.**

Telephone :01158231676, **Email:** msk-recruitment@nottingham.ac.uk

Thank you for taking the time to read this information sheet

11.7.4.5 Pages 1 & 2 of CAP-RA study questionnaire booklet

Dedicated to Outstanding care

Sherwood Forest Hospitals NHS Foundation Trust

University of Nottingham UK | CHINA | MALAYSIA

Central Aspects of Pain in Rheumatoid Arthritis (CAP-RA)

Baseline Questionnaire Booklet

We are developing a new questionnaire to measure how your central nervous system (brain and spinal cord) can be affected by, and can change pain and fatigue. We are investigating ways of identifying people with rheumatoid arthritis who might have or develop persistent pain or fatigue.

Please try to answer each question, even if it seems like some questions are repeated. Each question has been carefully selected and is equally important.

It is usually best to answer the questions with your first thoughts. Don't think too long and hard, there are no right or wrong answers and you may answer in any way you feel is best.

Staff from the research study are happy to help you with the questionnaire.

Thank you for your assistance with this area of research.

Chief Investigator – Professor David Walsh

IRAS reference No: 269143

Office use only

Centre ID:_____

Study ID: CAPRA

Date: _____ / _____ / _20_____

1. What is your date of birth?

☐☐ Day ☐☐ Month ☐☐☐☐ Year

2. What is your sex? ☐ Female ☐ Male

3. Choose one option that best describes your ethnic group or background?

☐ Asian ☐Black ☐White ☐Other

If other, please describe

4. What is your smoking status? (*A smoker is someone who smokes any tobacco product, either daily or occasionally.)*

☐Smoker ☐ Ex-smoker ☐Never Smoked

5. What is your employment status?

☐Employed/self-employed ☐Unemployed ☐Retired ☐Student

6. Has a doctor ever told you that you have any of these conditions or problems? ***Please select all that apply.***

☐Kidney Disease ☐ Hypertension ☐Fracture ☐Heart Attack

☐Depression ☐Diabetes ☐Cancer ☐Lung Disease

☐Other Cardiovascular Disease or Stroke ☐Ulcer or Stomach Problem

☐Other

If other, please describe

'Let us hear the conclusion of the whole matter: Fear God and keep his commandments: for this is the whole duty of man.'

Ecclesiastes 12:13 KJV