Understanding the patient experience of stiffness, and developing a stiffness patient-reported outcome measure in rheumatoid arthritis

Serena Elizabeth Halls

A thesis submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy

Faculty of Health and Applied Sciences
University of the West of England, Bristol

June 2016

Word count: 89651

Abstract

Rheumatoid arthritis (RA) is a chronic, systemic, inflammatory disease. Stiffness is a major symptom of RA which is commonly reported by patients, affects patients' daily life, and is relevant to patients in relation to fluctuating aspects of RA such as flare and low disease activity. Morning stiffness is also frequently used as an outcome measure both clinically and in research. Despite the relevance and uses of stiffness, it remains poorly understood and was omitted from the RA core set because of poor measurement properties. A pragmatic, mixed methods approach was used to better understand the patient experience of stiffness in people with RA and to develop and test a new RA stiffness patient reported outcome measure (PROM). It involved a systematic literature review, semi-structured interviews, focus groups, cognitive interviews, the development of appropriate candidate items to characterise stiffness and multivariate analysis of a survey using these items.

The systematic literature review found that current stiffness assessment is based on items that capture the duration or severity of morning stiffness. However, items were often poorly defined, highly variable in wording and format, had limited measurement property evidence and had not been developed according to current standards including collaboration with patients. Overall, there was no evidence regarding the most appropriate way to assess stiffness in RA, indicating the need for a new measure developed according to best practice PROM guidelines. Semi-structured interviews with RA patients provided an improved understanding of their experience of stiffness, demonstrated its relevance to patients and enabled the development of a conceptual model. These data also highlighted inconsistencies between current stiffness assessment and the patient perspective of this symptom. Focus groups with RA patients reinforced the stiffness conceptual model in a new sample, using a different method of data collection. They also provided information specifically addressing stiffness assessment from the patient perspective, including a number of concepts for measurement instrument development. These patient-driven concepts and qualitative data were tempered with measurement theory to develop a conceptually sound yet practically appropriate preliminary set of items for a new RA stiffness PROM. Preliminary items were reviewed and modified by RA patients in cognitive interviews. Following refinement, 45 candidate items (39 new items and 6 traditional stiffness items) were taken forward to a postal survey to develop and test the structure of a new RA stiffness PROM.

Analysis of the survey responses involved rigorous statistical testing including a series of iterative principal component analyses (undertaken initially with two different approaches), balancing Cronbach's alpha for internal consistency, bootstrapping for stability, and expert judgement for clinical appropriateness. The emergent structure was the Rheumatoid Arthritis Stiffness (RAST) questionnaire with 21 items in 3-components capturing 'stiffness severity', 'physical impact' and 'psychosocial impact'. The initial qualitative work enhanced its content validity and statistical testing for appropriate relationships with other measures of disease demonstrated good construct validity. These results provide support for RAST as an appropriate tool for use in future stiffness assessment.

The development of the RAST is important in recognising stiffness as a relevant patient symptom and is a significant step towards standardised stiffness assessment. Further testing in a fresh population will generate additional evidence of reliability and sensitivity to change to support its use. The RAST provides a measure for use in new investigations of disease mechanisms and response to therapy.

Acknowledgements

Thank you to all the patients who took part in this research and without whom this work would not have been possible. I am very grateful that you took the time to share your thoughts and opinions with me. Thank you also to the rheumatology teams at the BRI, North Bristol and Weston who facilitated recruitment.

I am very grateful to the University of the West of England who funded this research through a PhD Studentship and extended funding. I also appreciate the financial support provided to attend conferences to disseminate the results of this research.

To my fantastic supervisory team, I am incredibly grateful for your constant support and guidance. Professor Sarah Hewlett, I couldn't have asked for a better director of studies. Thank you for everything – you are an inspiration. Doctor Emma Dures, thank you for your tireless support and encouragement, and for your constant positivity. Doctor Jon Pollock, thank you for critical eye and for sharing your knowledge and expertise particularly with the statistical aspects of this thesis. Professor John Kirwan, thank you for always highlighting the things that no one else thinks of and for your endless enthusiasm. Thank you all for your thoughts, ideas and comments on every presentation, chapter and paper. I have learnt so much from you.

To patient research partners Gill Baker and Avis Edmunds, thank you for your hard work and support throughout this project.

Special thanks to Oonagh Wilson, Victoria Salmon and Katy Buchan. I don't quite know what I would have done without you all! I am very lucky to have gotten to know you and to have shared this journey with you.

Thank you also to the extended UWE team based at the Academic Rheumatology Department – you have been so helpful, friendly and supportive and I have really valued working with such a lovely group of people. In particular, thank you to Caroline Flurey and Julie Taylor for being constant sources of advice and guidance.

To my family, thank you for your constant love and support. Kath fy nghariad, diolch yn fawr iawn!

Table of contents

Abstract 2
Acknowledgements 3
Table of contents 4
List of boxes 9
List of figures 10
List of tables 12
List of abbreviations 16
Chapter 1: Rheumatoid arthritis and stiffness 18
1.1 Rheumatoid arthritis (RA) 18
1.2 Rheumatoid arthritis symptoms and consequences 19
1.2.1 Symptoms 19
1.2.2 Consequences 21
1.3 Management of rheumatoid arthritis 22
1.3.1 Non-pharmacological approaches 22
1.3.2 Pharmacological approaches 23
1.4 Measurement of disease activity 24
1.5 Relevance of stiffness 25
1.5.1 Historical relevance of stiffness 25
1.5.2 Relevance of stiffness to RA patients 26
1.5.3 Relevance of stiffness to healthcare professionals 29
1.5.4 Relevance of stiffness in research 31
1.5.5 Relevance of stiffness in disease activity 32
1.6 Current understanding of stiffness in RA 33
1.6.1 Stiffness pathophysiology 33
1.6.2 The patient perspective of stiffness in RA 33
1.7 Stiffness treatment and management 34
1.7.1 Pharmacological treatment and management 34
1.7.2 Non-pharmacological treatment and management 35
1.8 Stiffness in other rheumatic diseases 36
1.8.1 Polymyalgia rheumatica (PMR) 36
1.8.2 Ankylosing spondylitis (AS) 36
1.8.3 Osteoarthritis (OA) 37
1.9 Stiffness in other populations 38
1.9.1 Older populations 38

1.9.2 Stiff person syndrome 38
1.9.3 Phantom limb syndrome 38
Chapter 2: Stiffness assessment and PROM development 39
2.1 Stiffness assessment in RA 39
2.2 Patient reported outcome measures (PROMs) 40
2.2.1 What are PROMs? 40
2.2.2 Use of PROMs 41
2.3 PROM development theory 41
2.3.1 Concepts of measurement 41
2.3.2 Meeting the OMERACT filter 44
2.3.3 PROM development guidelines 44
2.3.4 Test theory 44
2.4 Current stiffness PROMs 45
2.4.1 Systematic literature review of stiffness PROMs 45
2.4.2 Identification of other stiffness measurement literature 64
2.5 Importance of appropriate stiffness assessment 72
Chapter 3: Purpose and structure of research 73
3.1 Purpose of research 73
3.1.1 Research aims 73
3.2 Researcher perspective 74
3.2.1 Prior knowledge 74
3.2.2 Epistemological position 74
3.3 Methodological approach 75
3.3.1 Mixed methods 75
3.3.2 Research design 75
3.3.3 Thesis structure 76
3.4 Research process and influences 78
3.4.1 Patient research partner involvement 78
3.4.2 Supervisory team 78
3.4.3 Reflexivity 79
Chapter 4: Understanding stiffness from the patient perspective (Study 1) 80
4.1 Background 80
4.2 Aims and objectives 81
4.3 Methods 81
4.3.1 Semi-structured interviews 81
4.3.2 Participant identification and sampling 81

4.3.3 Interview guide development 83
4.3.4 Interview procedure 84
4.4 Analysis....85
4.4.1 Analysis approach....85
4.5 Results....88
4.5.1 Participants....88
4.5.2 Thematic analysis....90
4.5.3 Stiffness descriptors and metaphors....112
4.6 Discussion....115
4.7 Conclusion....119
Chapter 5: Checking the conceptual model and investigating stiffness assessment (Study 2)....121
5.1 Background....121
5.2 Aims and objectives....121
5.3 Methods....122
5.3.1 Focus group methodology....122
5.3.2 Participant identification and sampling....123
5.3.3 Focus group topic guide development....124
5.3.4 Focus group procedure....126
5.4 Analysis....127
5.5 Results....128
5.5.1 Participants....128
5.5.2 Focus groups....130
5.6 Discussion....149
5.7 Conclusion....153
Chapter 6: Item development....154
6.1 Background....154
6.1.1 PROM development guidelines....155
6.2 Aims and objectives....155
6.3 Item development....155
6.3.1 Moving from the qualitative experience of stiffness to stiffness measurement....156
6.3.2 Process of item development....158
6.3.3 The final set of items for further testing....173
Chapter 7: Testing the draft content of the RA stiffness PROM with patients (Study 3)....174

7.1 Background 174
7.2 Aims and objectives 174
7.3 Methods 175
7.3.1 Cognitive theory 175
7.3.2 Cognitive interviewing methodology 175
7.3.3 Participant identification and sampling 176
7.3.4 Cognitive interview topic guide development 176
7.3.5 Cognitive interviewing procedure 177
7.4 Analysis 178
7.5 Results 178
7.5.1 Participants 178
7.5.2 Cognitive interviews 180
7.6 Discussion 213
7.7 Conclusions 216
Chapter 8: Developing the structure and content for an RA stiffness PROM (Study 4, part 1) 217
8.1 Background 217
8.1.1 Questionnaire development methodology 217
8.1.2 Multivariate analysis 218
8.1.3 The appropriateness of PCA 218
8.2 Aims and objectives 220
8.3 Methods 220
8.3.1 Postal survey 220
8.3.2 Patient identification and sampling 221
8.3.3 Questionnaire pack 222
8.4 Analysis and results 235
8.4.1 Stage 1 analysis: Data cleaning and descriptive statistics 235
8.4.2 Stage 1 results: Data cleaning and descriptive statistics 237
8.4.3 Stage 2 analysis: Identification of an appropriate analysis method 241
8.4.4 Stage 2 results: Identification of an appropriate analysis method 246
8.5 Discussion 262
8.5.1 Postal survey 262
8.5.2 Analysis approach 264
8.6 Conclusion 265
Chapter 9: Developing and testing the structure of an RA stiffness PROM (Study 4, part 2) 266

9.1 Background 266
9.1.1 Developing the provisional RA stiffness PROM 266
9.1.2 Concepts of measurement 266
9.2 Aims and objectives 267
9.3 Methods 267
9.4 Analysis and results 267
9.4.1 Stage 3 analysis: Item reduction and structure development 267
9.4.2 Stage 3 results: Item reduction and structure development 270
9.4.3 Stage 4 analysis: Validity testing 311
9.4.4 Stage 4 results: Validity testing 311
9.5 Discussion 317
9.6 Conclusion 322
Chapter 10: Summary and discussion 323
10.1 Thesis aims 323
10.2 Contributions to knowledge 323
10.2.1 Current stiffness assessment tools and measurement properties 324
10.2.2 Understanding stiffness 324
10.2.3 Development and content of a novel stiffness PROM (RAST) 326
10.2.4 Measurement property evidence for stiffness assessment tools 329
10.3 Implications for PRO methodology 332
10.4 Implications for research 333
10.5 Implications for clinical practice 335
10.6 Future research 337
10.6.1 Further development and validation of RAST 337
10.6.2 Specific or general nature of stiffness assessment 338
10.7 Strengths and limitations 338
10.8 Personal reflection 340
10.9 Thesis summary 341
References 342
Appendices 378
Appendix A: Data extraction form 379
Appendix B: Patient information sheet (Study 1) 384
Appendix C: Interview topic guide with prompts (Study 1) 386
Appendix D: Questionnaire pack 387
Appendix E: Study 1 coding tree 391
Appendix F: COREQ checklist (Study 1) 400

Appendix G: Patient information sheet (Study 2) 402
Appendix H: Original focus group topic guide (Study 2) 404
Appendix I: Revised focus group topic guide (Study 2) 405
Appendix J: Study 1 coding tree with Study 2 codes added 406
Appendix K: Study 2 coding tree 416
Appendix L: COREQ checklist (Study 2) 420
Appendix M: Identifying what is relevant to measure 422
Appendix N: Mapping the experience of stiffness to the measurement perspective 434
Appendix O: Early item development 442
Appendix P: Item tracking matrix 451
Appendix Q: Patient information sheet (Study 3) 468
Appendix R: Cognitive interview topic guide (Study 3) 470
Appendix S: Draft PROM items for cognitive interviews (Study 3) 471
Appendix T: Probes for cognitive interviews (Study 3) 483
Appendix U: COREQ checklist (Study 3) 487
Appendix V: Questionnaire pack (Study 4) 489
Appendix W: Frequency and distribution graphs (normality assessment where relevant) for non-stiffness items (Study 4) 506
Appendix X: Pearson's correlation coefficient (Study 4) 511
Appendix Y: Oblique promax rotation pattern and structure matrices (Study 4) 513
Appendix Z: Scoring the RAST 517
Appendix AA: Final layout of the RA stiffness PROM 519
Appendix BB: Abstract (Study 1) 522
Appendix CC: Abstract (Systematic literature review) 523
Appendix DD: Abstract (Study 4 development) 524
Appendix EE: Abstract (Study 4 testing) 525

List of boxes

Box 5.1: Questions presented to participants in focus group 1 127
Box 7.1: Cognitive interview instructions for participants 177
Box 7.2: Changes to item 1 181
Box 7.3: Changes to item 2 183
Box 7.4: Changes to item 3 185
Box 7.5: Changes to item 4 186
Box 7.6: Changes to items 6 and 7 188

Box 7.7: Changes to item 10....189
Box 7.8: Changes to item 11....190
Box 7.9: Changes to item 13....191
Box 7.10: Changes to item 16....192
Box 7.11: Changes to item 17....192
Box 7.12: Changes to item 19....193
Box 7.13: Changes to item 20....193
Box 7.14: Changes to item 22....194
Box 7.15: Changes to item 23....194
Box 7.16: Changes to item 24....195
Box 7.17: Changes to item 28....197
Box 7.18: Changes to item 30....197
Box 7.19: Changes to item 31....198
Box 7.20: Changes to item 35....199
Box 7.21: Refined impact triad items (39, 41, 44, 45)....205
Box 7.22: Changes to response shift items....210
Box 7.23: Changes to questionnaire introduction....211
Box 7.24: Additional items related to the introduction....212
Box 9.1: Rotated PCA and internal consistency analyses for item removal....275

List of figures

Figure 2.1: Flow diagram showing article selection process....47
Figure 3.1: PhD outline....77
Figure 4.1: Conceptual diagram of the patient experience of stiffness....90
Figure 6.1: Key areas for measurement from the conceptual framework....157
Figure 8.1: Recruitment flow diagram....237
Figure 8.3: Frequency and distribution graph....246
Figure 8.4: Frequency and distribution graph....246
Figure 8.5: Frequency and distribution graph....246
Figure 8.6: Frequency and distribution graph....246
Figure 8.7: Frequency and distribution graph....247
Figure 8.8: Frequency and distribution graph....247
Figure 8.9: Frequency and distribution graph....247
Figure 8.10: Frequency and distribution graph....247
Figure 8.11: Frequency and distribution graph....247
Figure 8.12: Frequency and distribution graph....247

Figure 8.13: Frequency and distribution graph....248
Figure 8.14: Frequency and distribution graph....248
Figure 8.15: Frequency and distribution graph....248
Figure 8.16: Frequency and distribution graph....248
Figure 8.17: Frequency and distribution graph....248
Figure 8.18: Frequency and distribution graph....248
Figure 8.19: Frequency and distribution graph....249
Figure 8.20: Frequency and distribution graph....249
Figure 8.21: Frequency and distribution graph....249
Figure 8.22: Frequency and distribution graph....249
Figure 8.23: Frequency and distribution graph....249
Figure 8.24: Frequency and distribution graph....249
Figure 8.25: Frequency and distribution graph....250
Figure 8.26: Frequency and distribution graph....250
Figure 8.27: Frequency and distribution graph....250
Figure 8.28: Frequency and distribution graph....250
Figure 8.29: Frequency and distribution graph....250
Figure 8.30: Frequency and distribution graph....250
Figure 8.31: Frequency and distribution graph....251
Figure 8.32: Frequency and distribution graph....251
Figure 8.33: Frequency and distribution graph....251
Figure 8.34: Frequency and distribution graph....251
Figure 8.35: Frequency and distribution graph....251
Figure 8.36: Frequency and distribution graph....251
Figure 8.37: Frequency and distribution graph....252
Figure 8.38: Frequency and distribution graph....252
Figure 8.39: Frequency and distribution graph....252
Figure 8.40: Frequency and distribution graph....252
Figure 8.41: Frequency and distribution graph....252
Figure 8.42: Frequency and distribution graph....252
Figure 8.43: Frequency and distribution graph....253
Figure 8.44: Frequency and distribution graph....253
Figure 8.45: Frequency and distribution graph....253
Figure 8.46: Frequency and distribution graph....253
Figure 8.47: Frequency and distribution graph....253
Figure 9.1: “Physical” sum score plotted against “Physical” regression factor score 305

Figure 9.2: “Severity” sum plotted against “Severity” regression factor score.....305
Figure 9.3: “Psychosocial” sum plotted against “Psychosocial” regression factor score.....305
Figure 9.4: Frequency and distribution for “Physical” 308
Figure 9.5: Frequency and distribution for “Severity”.....309
Figure 9.6: Frequency and distribution for “Psychosocial”.....309
Figure 9.7: Frequency and distribution for the sum of all components.....310

List of tables

Table 2.1: Different correlation cut-off definitions 43
Table 2.2: Description of current stiffness assessment measures.....49
Table 2.3: Correlations between items assessing different concepts 55
Table 2.4: Correlations between items assessing the same concept 56
Table 2.5: Correlations between stiffness items and disease activity 57
Table 2.6: Correlations between stiffness items and disability.....58
Table 2.7: Correlations between stiffness items and pain 59
Table 2.8: Papers that describe scale development and/or validation where stiffness item/s are included within the scale.....65
Table 3.1 Supervisory team characteristics.....79
Table 4.1: Interview guide.....83
Table 4.2: Participant demographic information 89
Table 4.3: Stiffness descriptors (number of transcripts identified in) 112
Table 5.1: Stiffness descriptors from Study 1 and the literature 125
Table 5.2: Areas of further discussion in focus groups 2 and 3 126
Table 5.3: Participant demographic information 129
Table 5.4: Comparison of participants across groups 130
Table 5.5: Stiffness descriptors generated by each focus group 137
Table 5.6: Patient preferences for stiffness descriptors.....138
Table 6.1: Traditional stiffness items.....164
Table 6.2: Reading age required for the preliminary stiffness items 173
Table 7.1: Participant demographic information 179
Table 8.1: Questionnaire pack contents, full item wording, abbreviated item wording and rationale.....223
Table 8.2: Demographic information across whole sample 238
Table 8.3: Participant demographic information 239
Table 8.4: Participant sociodemographic information.....240

Table 8.5: Draft items (n=45) and rationale for retention (n=38) or removal (n=7) based on response rate, distribution and correlations (assessing item suitability, Section 8.4.4.1) 256
Table 8.6: Eigenvalues, VAF, rotated component loadings, and communalities across PCA and NLPCA solutions 260
Table 9.1: Item loadings and Cronbach's alpha coefficient for component 1 ("Physical") 272
Table 9.2: Item loadings and Cronbach's alpha coefficient for component 2 ("Severity") 273
Table 9.3: Item loadings and Cronbach's alpha coefficient for component 3 ("Psychosocial") 273
Table 9.4: Item loadings and Cronbach's alpha coefficient for component 4 ("Location") 274
Table 9.5: Rotated PCA 1 275
Table 9.6: Rotated PCA 2 ('Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' (Draft item coped) removed) 276
Table 9.7: Rotated PCA 3 ('Over the past 7 days have you experienced RA stiffness all over?' (Draft item all over) removed) 277
Table 9.8: Rotated PCA 4 ('Has RA stiffness affected your sleep?' (Draft item sleep) removed) 278
Table 9.9: Rotated PCA 5 ('Have you experienced RA stiffness in your body (outside of your joints) over the past 7 days?' (Draft item in body) removed) 279
Table 9.10: Rotated PCA 6 ('Have your daily tasks and activities required more effort because of RA stiffness?' (Draft item effort) removed) 280
Table 9.11: Rotated PCA 7 ('Has your movement been restricted because of RA stiffness?' (Draft item movement) removed) 281
Table 9.12: Rotated PCA 8 ('Has RA stiffness made it difficult to do hobbies or activities you enjoy?' (Draft item hobbies) removed) 282
Table 9.14: Rotated PCA 10 ('Has RA stiffness made it difficult to open and close your fist?' (Draft item open/close fist) removed) 284
Table 9.15: Rotated PCA 11 ('Has RA stiffness reduced your strength to do tasks?' (Draft item strength) removed) 285
Table 9.16: Rotated PCA 12 ('Has RA stiffness made you slower (for example, unable to do things quickly)?' (Draft item slower) removed) 286
Table 9.17: Rotated PCA 13 ('Has RA stiffness made it difficult to get up after sitting for a while?' (Draft item get up after sitting) removed) 287

Table 9.18: Rotated PCA 14 (‘Have you had to concentrate to move your body because of RA stiffness?’ (Draft item concentrate) removed) 288
Table 9.19: Rotated PCA 15 (‘Has it taken you longer to do your daily tasks or activities because of RA stiffness?’ (Draft item take longer) 10.30 removed) 289
Table 9.20: Rotated PCA 16 (‘Have you needed help (from others or gadgets) because of RA stiffness?’ (Draft item need help) removed) 290
Table 9.21: Rotated PCA 17 (‘Have you had to work around your RA stiffness (or do things in a different way)?’ (Draft item work around) removed) 291
Table 9.22: Rotated PCA 18 (‘Over the past 7 days have you experienced RA stiffness after a period of immobility (for example, after sitting for a while)?’ (Draft item after immobility) removed).....292
Table 9.23: Rotated PCA 19 (‘Has RA stiffness made it difficult to balance without physically supporting yourself?’ (Draft item balance) removed).....293
Table 9.24: Rotated PCA 20 (‘Has RA stiffness made it difficult to do fine movements (for example, write with a pen)?’ (Draft item fine movement) removed) 294
Table 9.25: Rotated PCA 21 (‘Has RA stiffness made it difficult to grip or hold things?’ (Draft item grip) removed).....295
Table 9.26: Item loadings and Cronbach’s alpha coefficient for component 1 (“Physical”) 297
Table 9.27: Item loadings and Cronbach’s alpha coefficient for component 2 (“Severity”).....298
Table 9.28: Item loadings and Cronbach’s alpha coefficient for component 3 (“Psychosocial”).....298
Table 9.29: Rotated PCA (‘How would you describe the overall level of morning stiffness you have had from the time you wake up?’ (5-option ordinal scale) removed) 301
Table 9.30: Rotated PCA (‘Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week’) removed) 302
Table 9.31: Component structure and item placement during rotated PCA on 20 randomly selected subsets of the dataset 303
Table 9.32: Each component and the scale each item was scored on 304
Table 9.33: Correlation between simple summation and regression factor scores for each component 304
Table 9.34: Rescaled percentage score for “Physical” 306
Table 9.35: Rescaled percentage score for “Severity” 307

Table 9.36: Rescaled percentage score for “Psychosocial”.....307
Table 9.37: Correlation coefficients (*r*) and shared variance (R^2) between RAST (components and sum) and measures of disease.....313
Table 9.38: Correlation coefficients (*r*) and shared variance (R^2) between traditional stiffness items and clinical measures.....316

List of abbreviations

ACR	American College of Rheumatology
ADAS	Ankylosing Spondylitis Disease Activity Score
AIMS	Arthritis Impact Measurement Scales
AS	Ankylosing spondylitis
ASAS	Assessment of Spondyloarthritis International Society
ARA	American Rheumatism Association
AUSCAN	Australian/Canadian Osteoarthritis Hand Index
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BRAF	Bristol Rheumatoid Arthritis Fatigue Scales
BRI	Bristol Royal Infirmary
BSR	British Society for Rheumatology
CAPRA	Circadian administration of prednisone in rheumatoid arthritis
CASM	Cognitive aspects of survey methodology
CATPCA	Categorical principal component analysis
CBT	Cognitive behavioural therapy
CFA	Confirmatory factor analysis
CI	Confidence interval
COREQ	Consolidated criteria for reporting qualitative studies
CRP	C-reactive protein
CTT	Classical test theory
DAS	Disease activity score
DMARDs	Disease modifying anti rheumatic drugs
DR	Delayed-release
EFA	Exploratory factor analysis
EMS	Early morning stiffness
ES	Effect size
ESR	Erythrocyte sedimentation rate
ESSG	European Spondylarthropathy Study Group
EULAR	European League Against Rheumatism
FIQ	Fibromyalgia Impact Questionnaire
FIQR	Revised fibromyalgia Impact Questionnaire
FM	Fibromyalgia
GT	Grounded theory
HAQ	Health Assessment Questionnaire
ICF	International classification of functioning, disability and health
IL-1	Interlukin-1
IL-1ra	Interlukin-1 receptor antagonist
IL-6	Interlukin-6
IMD	Index of multiple deprivation
IPA	Interpretive phenomenological analysis
IQR	Interquartile range
IR	Immediate-release
IRT	Item response theory
ISPOR	International Society for Pharmaceconomics and Outcomes Research
KMO	Kaiser-Meyer-Olkin measure of sampling adequacy
LSOA	Lower layer super output areas
MHAQ	Modified Health Assessment Questionnaire
MRT	Modified-release tablet
MS	Morning stiffness

M-SACRAH	Modified Score for Assessment of Chronic Rheumatic Affections of the Hands
NBT	North Bristol Trust
NHP	Nottingham Health Profile
NLPCA	Nonlinear principal component analysis
NIC	Non-inflammatory complaint
NICE	National Institute for Health and Care Excellence
NRS	Numerical rating scale
NSAIDs	Non-steroidal anti-inflammatory drugs
OA	Osteoarthritis
OMERACT	Outcome Measures in Rheumatology
OR	Odds ratio
PAS	Patient Activity Scale
PCA	Principal component analysis
PDAS2	Patient-Based Disease Activity Score without ESR
PFQ	Preliminary Flare Questionnaire
PIS	Patient information sheet
PMR	Polymyalgia rheumatica
PtG	Patient global
PRO	Patient reported outcome
PROM	Patient reported outcome measure
PsA	Psoriatic arthritis
PV	Plasma viscosity
QAS	Questionnaire appraisal system
RA	Rheumatoid arthritis
RADAI	Rheumatoid Arthritis Disease Activity Index
RADAR	Rapid Assessment of Disease Activity in Rheumatology
RAID	Rheumatoid Arthritis Impact of Disease
RAPID	Routine Assessment of Patient Index Data
RAPS	Rheumatoid Arthritis Pain Scale
RAQoL	Rheumatoid Arthritis Quality of Life Questionnaire
RAST	Rheumatoid arthritis stiffness PROM
RCT	Randomised controlled trial
RE	Relative efficiency
REC	Research ethics committee
RIB	Receiving incapacity benefit
ROC	Receiver operator curve
SACMOT	Scientific Advisory Committee of the Medical Outcomes Trust
SACRAH	Score for Assessment of Chronic Rheumatic Affections of the Hands
SF-SACRAH	Short Form Score for Assessment of Chronic Rheumatic Affections of the Hands
SJC	Swollen joint count
SMD	Standardised mean differences
SMOG	Simple Measure of Gobbledygook
SRM	Standardised response mean
TJC	Tender joint count
TNF-α	Tumour necrosis factor alpha
TRT	Timed-release tablet
USDHHS FDA	US Department of Health and Human Services, Food and Drug Administration
VAF	Variance accounted for
VAS	Visual analogue scale
VS	Verbal scale
WOMAC	Western Ontario McMaster Osteoarthritis Index

Chapter 1: Rheumatoid arthritis and stiffness

This thesis aims to answer the research question 'What is the experience of stiffness in people with rheumatoid arthritis (RA) and how should it be assessed as a patient-reported outcome measure (PROM)?' This chapter will introduce RA, its symptoms, aetiology and treatment. It will also explore stiffness specifically within the context of RA, and more broadly within other populations.

1.1 Rheumatoid arthritis (RA)

RA is a chronic, systemic, inflammatory condition causing synovitis and resulting in pain, swelling and stiffness (Arthur and Hill, 2006). Typically, RA affects the small joints in the hands, feet and wrists but can affect any synovial joint (Hakim, Clunie and Haq, 2011). Individuals with RA typically experience fluctuations of disease activity between high (flare) and low states, and now with modern treatments, remission (Smolen *et al,* 2010).

The American Rheumatism Association (ARA) defined a set of classification criteria. Four of the seven criteria must be met for a classification of RA, the first four of which must have been present for at least six weeks: morning stiffness lasting over one hour; soft tissue swelling of at least three joints; swelling of hand or wrist joints; symmetrical swellings; rheumatoid nodules; positive rheumatoid factor; and radiographic changes (Arnett *et al,* 1988). More recently, the 2010 American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) classification criteria (Aletaha *et al,* 2010) were developed, consisting of weighted scores from four categories; joint symptoms (number of small or large joints); serology (presence and level of rheumatoid factor and/or anti-citrullinated protein antibody); acute-phase reactants (normal or abnormal C-reactive protein (CRP)/erythrocyte sedimentation rate (ESR)); and symptom duration (<6 weeks or >6 weeks) (Aletaha *et al,* 2010). These criteria focus less on radiographic damage, symmetry and rheumatoid nodules, which may not be present in early RA (Kay and Upchurch, 2012).

RA is prevalent worldwide and can affect people of any age and gender, however it is three times more common in females (National Institute for Health and Care Excellence (NICE), 2009). RA affects around 400,000 people in the UK (NICE, 2009), approximately 1% of the population (Symmons *et al,* 2002). Although its aetiology is unknown, smoking (Sugiyama *et al,* 2010), genetics (Silman and Pearson, 2002), and obesity (Symmons *et al,* 1997) are risk factors. The annual combined indirect and

direct cost of RA has been estimated as between £3.8 and £4.8 billion in the UK (NICE, 2009). The outcomes for individuals with RA include reduced function, reduced life expectancy (Naz and Symmons, 2007), and increased work disability (Burton *et al,* 2006; Sokka *et al,* 2010). Higher rates of psychological outcomes such as depression (Dickens *et al,* 2002) and anxiety (Isik *et al,* 2007) are also reported in RA populations than in the general population. However, outcomes have improved and during the past two decades, RA patients have been shown to be less psychologically distressed and physically disabled. This has been suggested to be a result of improved treatment and management including earlier diagnosis, increased emphasis on physical activity, and more effective medications (Overman *et al,* 2014), in which there have been substantial changes with the introduction of new medications and combination regimens (Smolen *et al,* 2010). These developments in medications such as biologics (Section 1.3.2) aim to target the actions of pro-inflammatory cytokines (White and Bryer, 2006). Cytokines are proteins that transport messages to different body systems, and pro-inflammatory cytokines specifically increase inflammation as part of the immune response (Oliver, 2006). Cytokines also promote joint destruction, and are thought to play a role in aspects such as fatigue within the RA process (Choy, 2012). Particularly relevant in RA are pro-inflammatory cytokines interleukin 6 (IL-6) and tumour necrosis factor alpha (TNF-*α*) (McInnes and Schett, 2003). Improved understanding of these cytokines and their role at a cellular level have enabled the development of medications that work by disrupting the production of pro-inflammatory cytokines (Oliver, 2006). Cytokines follow circadian rhythms, demonstrating a peak in the night and early morning (Cutolo *et al,* 2003). It is suggested that variation in symptoms including stiffness relates to the variation in cytokines such as IL-6 (Perry *et al,* 2009).

1.2 Rheumatoid arthritis symptoms and consequences

1.2.1 Symptoms

1.2.1.1 Pain

Pain is a common and complex symptom with the purpose of protection (Harvey, 1987). There are different theories explaining the cause of pain, such as gate control theory (Melzack and Wall, 1965). Pain has been defined as "[…] whatever the experiencing person says it is, existing whenever he says it does" (McCaffrey, 1983, cited in Hill, 2006, p.218) to reinforce its subjective nature. It is managed using pharmacological interventions such as analgesics or steroids, and therapies such as heat and cold application (Hill, 2006). Pain has been linked to other symptoms

including disability (Sprangers *et al,* 2000) and stiffness (e.g. Rhind, Unsworth and Haslock, 1987; Hazes, Hayton, and Silman, 1993).

In a study involving the 12 dimensions of health status in the Arthritis Impact Measurement Scales (AIMS2) (Meenan *et al,* 1992), 58 female participants identified pain as the most relevant dimension of impairment (Minnock, FitzGerald, and Bresnihan, 2003). In another survey of 1,024 male and female RA patients, pain was identified by 68.6% (n=702) as the preferred area for health improvement (Heiberg and Kvien, 2002). However, both studies were based on the AIMS2, which although addresses patient preferences, its dimensions were not patient generated (Meenan *et al,* 1992). Therefore, when the study by Minnock *et al.* (2003) was repeated with the same patients in a follow-up study with the addition of fatigue, an important patient generated symptom (Hewlett *et al,* 2005a), 65% of women prioritised fatigue for improvement over pain (Minnock and Bresnihan, 2004; 2008).

1.2.1.2 Fatigue

Fatigue is another commonly experienced RA symptom (Wolfe, Hawley, and Wilson, 1996; Overman *et al,* 2015), that is important to patients, yet has been reported as being ignored by healthcare professionals (Hewlett *et al,* 2005b). However, the importance of fatigue has recently been recognised, leading to the recommendation that it is assessed in all RA clinical trials (Kirwan *et al,* 2007) and the development and validation of the Bristol RA Fatigue (BRAF) scales (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b; Dures *et al,* 2013). There is still limited understanding of the underlying cause of fatigue. A recent systematic review identified 25 studies relating to fatigue causes. Within these, RA related aspects, physical function, cognitive and emotional function, social and environmental aspects, and female sex were identified as possible causes of fatigue however, these findings were inconsistent across studies (Nikolaus *et al,* 2013). A conceptual model of RA fatigue has been proposed which suggests interactions between three key areas including RA disease processes (e.g. inflammation), cognitive and behavioural aspects (e.g. illness beliefs), and personal life aspects (e.g. responsibilities) (Hewlett *et al,* 2011a). Despite the limited understanding of fatigue, interventions have demonstrated fatigue reduction indirectly by improving aspects identified in the conceptual model (Hewlett *et al,* 2011b). For example, a fatigue self-management intervention has been shown to be effective at reducing RA fatigue (Hewlett *et al,* 2011b; Hewlett *et al,* 2014), based on cognitive behavioural therapy (CBT) which targets the links between thoughts, feelings, and behaviours to encourage change (Sage, 2008).

1.2.1.3 Stiffness

For individuals with RA, stiffness is a major problem (Hill, 2006). As this symptom is the focus of this research, its relevance including its emergence and use in different contexts is discussed in detail in Section 1.5.

1.2.2 Consequences

1.2.2.1 Psychological consequences

Diagnosis of chronic conditions can result in psychological consequences such as depressive symptoms (Polsky *et al,* 2005). In rheumatic diseases, the prevalence of anxiety and depression are twice that found in the general population (Geenan *et al,* 2012). A study investigating comorbidities in RA patients (n=7818) on the British Society for Rheumatology (BSR) Biologics Register reported that depression was one of the most frequent conditions (n=1491, 19%) (Hyrich *et al,* 2006). As such, psychological support provision is an important part of patient care in UK (Luqmani et al, 2006; Luqmani et al, 2009) and international guidelines (American College of Rheumatology, 2002). Consistent with the above, a recent survey of the psychological support preferences of patients with inflammatory arthritis (n=1210) indicated that demand for psychological support was high, with only 6% of participants reporting that social and emotional aspects were not relevant. Yet despite demand, only 23% of participants reported being asked about such issues by members of their clinical team, indicating that current provision of psychological support does not meet patient need (Dures *et al,* 2014).

1.2.2.2 Disability

Disability has been defined by the World Health Organisation as “an umbrella term, including impairments (a problem in body function or structure), activity limitation (difficulty executing tasks or actions), and participation restriction (difficulty with involvement in life situations)” (World Health Organisation, 2015). In an RA population, disability is often assessed using the disability component of the Health Assessment Questionnaire (HAQ) (Fries *et al,* 1980). Disability in RA increases with disease duration (Scott *et al,* 2000) and higher amounts of joint damage are related to increases in disability over time (Bombardier *et al,* 2012). It is thought that uncontrolled inflammatory activity causes joint damage which subsequently leads to disability, although this relationship is not fully understood (Scott *et al,* 2000). However, with the recent improvements in medications, and better control of inflammation, disability in RA has declined (Krishnan *et al,* 2011). This was

demonstrated in a prospective cohort study involving 4651 patients from a USA and Canadian database which demonstrated annual mean disability reductions from 1993 to 2006. Use of biologic therapies was introduced in 1998 thus providing a likely explanation for this reduction. However, this does not explain the declines in disability seen before this time (between 1993 and 1998) which the authors suggest may be attributed to the use of medications such as methotrexate, patient selection or decline in the severity of RA over time (Krishnan *et al,* 2011).

A key area of disability in an RA context is work disability. Early research into work disability reported that 50% of people with RA stopped work within a decade after diagnosis (Yelin, Henke and Epstein, 1987). It was anticipated that the improvement in medications especially the introduction of biologics, would result in less work disability (Verstappen, Jacobs and Hyrich, 2007). However, work disability still appears problematic. One cohort study reported that of the 8082 participants who were employed at the onset of RA, 43.8% were not working 12.8 years after onset. Of these, 22.7% (n=1837) defined themselves as disabled, 30.5% (n=2496) had stopped work for health reasons, and 20.6% were receiving benefits as a result of disability (n=1236) (Wolfe, Allaire and Michaud, 2007). Another cohort study across 32 countries involving 8039 patients reported that of the 5493 participants <65 and working at symptom onset, 37% reported work disability as a result of RA (Sokka *et al,* 2010). Work disability appears to remain an important issue in RA however, its assessment and evaluation is difficult given different definitions, social policies, and cultural attitudes.

1.3 Management of rheumatoid arthritis

The management of RA involves symptom relief, function maintenance and modification of disease activity, specifically the prevention of erosive damage and the achievement of remission (Cornell, 2007). Current management for RA is based on a combination of non-pharmacological and pharmacological approaches.

1.3.1 Non-pharmacological approaches

A multidisciplinary team approach to RA management including rheumatologists, nurses, physiotherapists, occupational therapists, podiatrists, social services, and surgery is recommended by the NICE (2009) and the BSR (Luqmani *et al,* 2006; Luqmani *et al,* 2009). Randomised controlled trials (RCTs) have demonstrated the effectiveness of these in RA for example nurse-led care (Ndosi *et al,* 2014) and

occupational therapy (Hammond, Young and Kidao, 2004). In addition to multidisciplinary team management, patients may take complementary therapies although further research is required to demonstrate the benefits and/or risks of these therapies (Cornell, 2007; Arthritis Research UK, 2012; Arthritis Research UK, 2013a).

1.3.2 Pharmacological approaches

Pharmacological approaches to RA management rely on combinations of different medications. First-line therapies include analgesics (e.g. paracetamol), non-steroidal anti-inflammatory drugs (NSAIDs), and glucocorticoids (e.g. prednisolone) which focus on the control of symptoms. Second-line therapies include disease modifying anti-rheumatic drugs (DMARDs, e.g. methotrexate) which target the underlying disease process, and focus on suppression of disease activity. A recent development in the pharmacological treatment of RA is biologic therapies which are genetically engineered treatments (White and Bryer, 2006). Biologics or targeted therapies, work by blocking or altering the actions of pro-inflammatory cytokines and specifically target the cell or process that causes inflammation (White and Bryer, 2006). Currently there are three targets of biologics; interlukin-1 (IL-1), TNF-*α*, and CD20 B cells. Interlukin-1 receptor antagonist (IL-1ra), in the form of licenced therapy Anakinra, works by altering the action of IL-1 (Mertens and Singh, 2009) by binding to IL-1 receptors and maintaining an appropriate inflammatory response (Oliver, 2006). Similarly, the actions of TNF-*α* can be mediated by anti-TNF-*α* therapies. There are different types of anti-TNF-*α* including Adalimumab, Etanercept, and Infliximab which have slightly different actions e.g. Infliximab binds to TNF-*α* and stops it functioning while Etanercept stops TNF-*α* binding to its receptor (White and Bryer, 2006). Finally, the target of Rituximab is to reduce the number of B cells (Lopez-Olivo *et al,* 2015), thus reducing the production of auto-antibodies e.g. rheumatoid factor (White and Bryer, 2006). Cochrane reviews have been performed demonstrating the effectiveness of second-line treatments in RA (e.g. Navarro-Sarabia *et al,* 2005; Mertens and Singh, 2009; Lopez-Olivo *et al,* 2015).

For individuals with newly diagnosed RA, a combination of DMARDs and glucocorticoids (oral, intramuscular or intra-articular) are recommended (NICE, 2009). As biologics are considerably more expensive than other treatments, patients must meet specific criteria, such as a minimum disease activity score (Section 1.4) or intolerance to or inefficacy of other medications, before being considered for these treatments (NICE, 2009). However, despite the recent developments in treatment

such as biologic therapies and the effectiveness of the treatments available, there is no cure for RA.

1.4 Measurement of disease activity

In order to monitor each patient and ensure that the management of their disease is effective, a standardised measure of disease activity is used. The disease activity score (DAS) (van der Heijde *et al,* 1990) is a composite score including a tender joint count (TJC), a swollen joint count (SJC), ESR, and a patient global (PtG) assessment of general health on a visual analogue scale (VAS). The DAS28 is more concise and focuses on the assessment of 28 joints (Prevoo *et al,* 1995). DAS and DAS28 scores are calculated using a weighted formula producing final scores ranging between 0-9.4, from which, disease activity can be classified as low (≤3.2), moderate (3.2-5.1), or high (>5.1) (Prevoo *et al,* 1995; Fransen, Stucki and van Riel, 2003). There are also different versions of the DAS28 for example, using CRP rather than ESR (Fransen, Stucki and van Riel, 2003). The DAS28 is a well validated tool that is useful for monitoring disease activity clinically and in research (van Riel and Schumacher, 2001; Fransen, Stucki and van Riel, 2003). However, there are a number of criticisms of the DAS28. Firstly, it does not include the feet and in early RA patients with mainly foot involvement this may underestimate scores (Bakker *et al,* 2011). Despite this, joint counts that do not include the feet have been shown to discriminate between patients with different levels of disease activity, and demonstrate as much validity as joint counts that do include them (Prevoo *et al,* 1995). It has also been shown that despite foot involvement being common, measurement precision is not improved by including this information in joint counts (Siemons *et al,* 2013). In addition, there are practical (de Souza, Williams and Lempp, 2016) and accuracy (van Tuyl *et al,* 2011) related reasons for exclusion of the feet. Therefore, this remains a debated topic in the literature (Siemons *et al,* 2013). Another criticism is that three of the four DAS28 measures are assessed by a clinician which conflicts with the recent increased focus on the patient's perspective, especially in outcome assessment (Kirwan *et al,* 2003). Furthermore, the weighting given to the PtG VAS is the smallest of all the DAS28 components (0.014xPtG) and the DAS28 can be scored without the PtG VAS (Prevoo *et al,* 1995). Development of the DAS was based on the clinical judgement of six rheumatologists (van der Heije *et al,* 1993) and did not involve patients including during the wording of the PtG VAS. The wording of the PtG VAS is not standardised, and a systematic review of the reporting of patient reported outcomes (PROs) found that PtG VAS wording differs across studies (Kalyoncu *et al,* 2009). In a study of five

different versions of the PtG VAS, different versions were shown to produce different DAS28 scores (French *et al,* 2013).

1.5 Relevance of stiffness

Stiffness is a major problem for people with RA (Hill, 2006). To understand the current stiffness knowledge base, comprehensive searches of electronic databases (Medline (via EBSCO), Allied and complementary medicine (AHMED), Cumulative index to nursing and allied heath (CINAHL plus), PsychINFO) were performed to identify research published up to December 2015. In addition to database searches, evaluation of known articles, theses and reviews, and their reference lists, expert suggestions, and other hand searching was performed. Searches for keywords included three key topic areas: 1) stiffness (e.g. stiffness, morning stiffness, early morning stiffness, joint stiffness); 2) population (e.g. RA, inflammatory arthritis); 3) instrument (e.g. patient report, PROM, questionnaire). The not function and title field selection were often used for specificity (e.g. not arterial stiffness, vascular stiffness, aortic stiffness).

1.5.1 Historical relevance of stiffness

Traditionally this symptom has been termed early morning stiffness (EMS), but also referred to as morning stiffness (MS), or stiffness. As will be discussed later (Section 1.6), there is little evidence of the appropriate term to use when discussing stiffness. Therefore, this thesis will use the term stiffness unless otherwise specified by referenced papers. In early RA literature it was stated that “morning stiffness is an almost universal manifestation of active rheumatoid arthritis” (Lansbury, 1956, p.11). Duration of MS was included in early disease activity composite assessments such as the Lansbury index (Lansbury, 1958), the Mallya-Mace index (Mallya and Mace, 1981), the Stoke index (Jones *et al,* 1993) and the Paulus Criteria (Paulus *et al,* 1990). Later the presence of “morning stiffness in and around the joints lasting at least 1 hour before maximal improvement” was included in the 1987 ACR classification of RA (Arnett *et al,* 1988, p.315). Furthermore, absence or short duration (<15 minutes) of MS was defined as one of six preliminary remission criteria (Pinals *et al,* 1981). The inclusion of stiffness as part of key RA criteria reflected stiffness being a common feature of RA that was frequently reported by patients (Scott, 1960; Vliet Vlieland *et al,* 1997) and considered an indicator of inflammatory activity (Lansbury, 1956). However, when the classification criteria were updated (Aletaha *et al,* 2010), although MS duration was considered, it was not included because it did not have enough

predictive value for DMARD initiation (Kay and Upchurch, 2012). Stiffness was also omitted from the ACR core set for assessing disease activity because it was not sensitive to change (Felson *et al,* 1993). As the revised remission criteria were based on the ACR core set, stiffness was not included (Felson *et al,* 2011). Furthermore, during the development of the DAS (van der Heije *et al,* 1990; van der Heije *et al,* 1992), although duration of MS was considered as one of the 19 tested variables, it was not included in the final score (van der Heije *et al,* 1990; van der Heije *et al,* 1992). Despite the DAS and the Mallya index (which includes MS duration) demonstrating the best validity in the assessment of disease activity, duration of MS was among a number of variables that performed poorly as individual variables (van der Heije *et al,* 1992). The usefulness of using MS as a single variable in the assessment of disease activity was also questioned (van der Heije *et al,* 1992). Therefore currently, there is no obligation in clinical or in research settings to routinely collect information about stiffness, although more recently this has been challenged with the inclusion of stiffness in the core set to assess flare (Bykerk *et al,* 2014a) (Section 1.5.2.3).

1.5.2 Relevance of stiffness to RA patients

Despite not being consistently collected or reported, there is a growing body of research highlighting the relevance of stiffness to people with RA. Recent studies have demonstrated that stiffness has considerable impact on patients' daily lives in a number of domains.

1.5.2.1 Work and disability

RA is associated with work disability (Burton and Lloyd, 2006; Sokka *et al,* 2010) and stiffness appears to have a specific role within that. In an observational study of RA patients exploring the relationship between MS and early retirement, MS was shown to be important (Westhoff *et al,* 2008). Of the 1023 patients in the cohort at baseline, 389 (38%) were under the retirement age cut-off of 61 years and not fulfilling other roles such as homemaker and therefore were at risk of early retirement. At the three year follow up, 65 (17%) of the 389 had taken early retirement specifically due to RA and within that population, early retirement was three times more likely in patients with severe MS, compared to those with mild MS (Westhoff *et al,* 2008). There has also been work conducted into the effect of MS on individuals who remain in employment. A survey across 11 European countries explored RA patients' perceptions of the impact of MS on work life (Mattila, Buttgereit and Tuominen, 2014).

Of the 1061 participants, 534 (50%) were currently in work (self-employed or full/part-time paid work) although the nature of the work was not included. Of these, 50 (15%) reported taking sick leave in the past month as a direct result of MS, 176 (33%) reported arriving late at work because of MS, and 250 (47%) felt that MS had an adverse effect on their work performance every week (Mattila, Buttgereit and Tuominen, 2014). Although demonstrating marked impact of MS on work, the authors highlighted that study participants were selected on experiencing MS on at least three days per week. This may have introduced bias relating to the impact of MS and has implications for the relevance of findings to the wider RA population, although given that stiffness is a common feature of RA (e.g. Khan *et al,* 2009), this may not be problematic. Given the importance to patients of remaining in employment (Grønning, Rødevand and Steinsbekk, 2010), and the broader relevance of reducing work disability in people with RA, this is an area requiring further investigation.

1.5.2.2 Quality of life

A large explorative survey investigated the impact of impaired morning function on RA patients' quality of life (da Silva, Phillips and Buttgereit, 2011; Phillips and Dow, 2012). Impaired morning function was reported to have a significant impact on responders' quality of life (82%) and work life (73%). Morning activities such as getting out of bed and dressing were also affected, with between 50-72% of patients being unable to carry out these tasks unimpaired. This affected participant's emotional state with respondents indicating that difficulties completing their morning activities made them feel frustrated (58%), angry (32%) and drained (14%) (da Silva, Phillips and Buttgereit, 2011). In further analysis, 84% indicated that impaired morning function had a significant effect on quality of life, which increased with MS severity (Phillips and Dow, 2012). These are important findings that demonstrate the significant impact of stiffness on RA patients' quality of life. However, it is difficult to determine how much of that impaired morning function was specifically due to stiffness (rather than pain or disability) given the focus on morning function. Yet, this work demonstrates the importance of understanding the specific effect of stiffness and highlights the relevance of stiffness to RA patients.

Another study investigated what RA patients would be willing to pay for reduction in or elimination of MS (Tuominen, Tuominen and Möttönen, 2011; Tuominen, Tuominen and Möttönen, 2012). Although three different approaches to estimating the monetary value were used producing quite different results, all approaches consistently identified that the impact of MS was between five and eight fold higher

for patients with severe MS (60-100 on numerical rating scale (NRS)) compared with patients with mild MS (0-29 on NRS). When only considering the willingness-to-pay estimates, patients with longer duration MS (≥60 minutes) reported that they would pay more than patients with shorter duration MS (10-59 minutes or <10 minutes) for every percentage reduction (25%, 50%, 75%, 100%). This was mirrored for MS severity where patients with more severe MS (60-100 on NRS) reported that they would pay more than patients with less severe MS (mild=0-29, moderate=30-59). Interestingly, patients with severe MS were willing to pay more than patients experiencing ≥60 minutes MS per day for a 100% reduction in the symptom (€47.9/£33.71 vs. €21.7/£15.27) (Tuominen, Tuominen and Möttönen, 2011; Tuominen, Tuominen and Möttönen, 2012). It is not known whether the same participants reported comparable levels for both severity and duration (e.g. whether participants reporting high severity are the same as those reporting long duration). If reported by different participants, this may explain differences in monetary value, although differences in the impact of MS severity and MS duration have also been reported in other work (Mattila, Buttgereit and Tuominen, 2014) (Section 2.4.1.4). These results could also have been influenced by the phrasing of the questions used as MS severity was reported on a NRS between 0 (best possible situation) and 100 (worst possible situation) which does not fit entirely with the categorisations of mild, moderate and severe.

1.5.2.3 RA disease activity

Stiffness is also important to patients in relation to RA disease activity. RA disease activity fluctuates meaning that patients typically experience periods of flare and low disease activity. Stiffness has been reported to be experienced by RA patients in different disease states. In a recent cohort study involving 5439 RA patients from the QUEST-RA database, it was reported that stiffness was experienced by 2884 (79%) of those patients with active disease compared to 614 (41%) patients with less active disease (Khan *et al,* 2009). Definitions of disease activity were consistent with standard classifications (Prevoo *et al,* 1995; Fransen, Stucki and van Riel, 2003). In contrast, in a smaller interview-based study involving 93 RA patients, stiffness was reported by 43 (89%) patients with active disease and 35 (81%) patients with inactive disease (Hazes, Hayton, and Silman, 1993). In this case however, the referring physician classified each patients' disease as active or inactive providing a more subjective classification than the DAS28. However, both studies demonstrate that stiffness appears to be commonly experienced across disease activity states.

In relation to remission, a recent study to understand the patient perspective of remission was performed (van Tuyl *et al,* 2015). This study employed qualitative focus groups in three centres across Europe and included 47 participants in total, all of whom either met the ACR/EULAR remission criteria (Felson *et al,* 2011), were self-defined as currently being in remission, or had past experience of being in remission. Three main themes were identified: symptoms, impact and normality. Within symptoms, patients highlighted stiffness reduction as crucial before they would consider themselves to be in remission (van Tuyl *et al,* 2015).

In terms of flare, qualitative research by the Outcome Measures in Rheumatology (OMERACT) flare working group has indicated that MS is an important influence on patients' decisions to seek medication review (Hewlett *et al,* 2012). An international patient and professional Delphi exercise then aimed to identify the relevant domains in the assessment of flare. Stiffness was prioritised as a potential item for a core set of flare measures (79% consensus) and patients and healthcare professionals were equally likely to classify stiffness as an essential characteristic (80% and 76%, $p>0.05$) (Bartlett *et al,* 2012). Having identified potential items for a flare core set, at OMERACT 11 these domains were discussed and assessment of each was considered. In the final consensus vote, 91% of delegates agreed that stiffness should be included as a core domain to assess RA flare and the identification or development of stiffness assessment methods was part of the resultant research agenda (Bykerk *et al,* 2014a). This discussion was continued in breakout groups at OMERACT 12 (Bartlett *et al,* 2015, Orbai *et al,* 2015).

1.5.3 Relevance of stiffness to healthcare professionals

Stiffness appears relevant to both patients and healthcare professionals when assessing remission and high disease activity (van Tuyl *et al,* 2015; Bartlett *et al,* 2012). A survey (Section 1.5.2.2) involving 518 rheumatologists and 750 RA patients investigated the impact of impaired morning function on RA patients (da Silva, Phillips and Buttgereit, 2011). In the study, patients and rheumatologists completed group specific questionnaires on the same topic. Patients reported a mean duration of MS and pain of 83 minutes which was slightly longer than that perceived by rheumatologists (70 minutes). Similarly, when considering the duration of morning function impairment, rheumatologist perceptions were consistent with patient reports. 95% of rheumatologists considered that impaired morning function had a significant impact of patients' quality of life (da Silva, Phillips and Buttgereit, 2011). However, these results only contain the views of rheumatologists and it may be that other

groups of health professionals have differing opinions. Although not reported in the paper directly but in a review by one of the principal authors (Buttgereit, 2011), 38% of rheumatologists indicated that patients should accept impaired morning function as a result of RA. This indicates that despite demonstrating understanding of the impact of impaired morning function on patients, some rheumatologists may not perceive this as an aspect requiring intervention or treatment. Further research with this group of health professionals demonstrated that rheumatologist's use patients' self-reports of stiffness as a crucial variable in decision-making for changing medication (Soubrier *et al,* 2006). This study aimed to identify the variables that were important in predicting change in DMARDs by reviewing demographics, disease activity, and outcomes of 204 consecutive patients attending outpatients. Rheumatologists were blinded to the intention of the study to avoid bias. MS duration was identified as independently associated with decisions to initiate DMARD change. It was the second highest influence on judgments (standardised odds ratio (OR) 3.38), exceeded only by SJC (standardised OR 5.24), and higher than both pain (standardised OR 0.98) and CRP (standardised OR 2.8) among other variables, which was similar to results reported in a previous study involving rheumatologists (Kirwan *et al,* 1984). The study suggested that MS was so relevant in DMARD change initiation because it is an indicator of inflammatory activity, despite not being included in the DAS28 or ACR core set (Soubrier *et al,* 2006).

In relation to other professional groups, a survey explored the management practices of Dutch physical therapists (n=233) when managing people with RA. The study reported that 168 (72%) of physical therapist responders 'always', 59 (25%) 'sometimes', and six (3%) 'never' assess MS (Hurkmans *et al,* 2012), demonstrating the use of stiffness in clinical assessment in this population. However, although it appears that Dutch physical therapist practices are in accordance with Dutch physiotherapy guidance (Hurkmans *et al,* 2011), there is no evidence of how this translates to other professional groups or practice in different countries. There was also no indication as to the way in which stiffness was assessed. In contrast, a small survey involving 32 experts in RA and/or spondylarthropathy were asked to complete a questionnaire about the 1987 RA classification criteria and the 1991 European Spondylarthropathy Study Group (ESSG) criteria for spondylarthropathy (Berthelot *et al,* 2002). The removal of MS was suggested in seven (22%) of the returned questionnaires. However, within the questionnaire there was a specific question about the removal of one or more clinical criteria with a dichotomous yes/no response

option and MS was given as an example. This may have biased responses towards this outcome.

1.5.4 Relevance of stiffness in research

Despite being commonly experienced by patients, until recently stiffness has rarely been the focus of research studies. As described earlier (Section 1.5.2.2), research has explored impaired morning function (e.g. da Silva, Phillips and Buttgereit, 2011; Phillips and Dow, 2012), a term defined as "stiffness and pain in the joints (particularly the hands) first thing in the morning that causes reduced strength, grip and mobility and results in difficulty to function or perform tasks" (da Silva, Phillips and Buttgereit, 2011, p.7). However, this definition makes it difficult to differentiate stiffness from pain and disability. This is made more difficult given the relationship between pain and stiffness (Rhind, Unsworth, Haslock, 1987; Lineker *et al,* 1999). One study explored the patient definition of stiffness with and without the use of a list of descriptors in three categories (difficulty of movement (n=7), pain (n=4), and sensations (n=2)) which was developed with patient input. Patient-generated descriptions were classified under the three categories and difficulty of movement followed by pain were most frequently used. As it was common for patients to provide combinations of descriptors, it was suggested that the patient definition of stiffness was an inter-relationship between pain and limited movement (Rhind, Unsworth, Haslock, 1987). These findings were consistent with a qualitative study in which some RA patients related MS to pain, particularly when discussing flare (Lineker *et al,* 1999). However, in a study investigating word meanings in relation to symptoms involving patients with rheumatic conditions (RA (n=100), fibromyalgia (FM) (n=50), ankylosing spondylitis (AS) (n=50), and osteoarthritis (OA) (n=50)), it was reported that patients were able to distinguish pain from stiffness (Helliwell, 1995). External to the patient perspective of the relationship between pain and stiffness, early work suggested that pain and stiffness are likely the result of different mechanisms (Ingpen, 1968). However, this position has seen little development given the limited understanding of the pathophysiological causes of stiffness (Section 1.6.1).

There has been a decline in the reporting of stiffness in clinical trials (Labitigan *et al,* 2010). In a study comparing the characteristics and outcomes reported of RA patients who participated in RCTs in the 1980s and 2000s, it was found that the reporting of MS decreased between the time periods. The study identified 114 and 172 RCTs from the two time points respectively and within those studies, MS was reported in 51% of in the studies in the 1980s but in only 27% of studies in 2000s (Labitigan *et*

al, 2010). This figure is consistent with a systematic review that was performed between 2005 and 2007 to explore the use of PROs in RA trials. The review identified 109 papers and found that MS was reported in 27% (n=29) of these (Kalyoncu *et al,* 2009). MS, function, pain, and PtG were the only PROs reported in >25% of articles (Kalyoncu *et al,* 2009).

1.5.5 Relevance of stiffness in disease activity

Stiffness is considered an indicator of inflammatory activity in RA (e.g. Lansbury, 1956; Hazes *et al,* 1994; Soubrier *et al,* 2006) and was included in many early disease activity composite assessments (e.g. the Lansbury index, Lansbury, 1956). However, it is not included in the DAS or DAS28 (van der Heije *et al,* 1990; Prevoo *et al,* 1995) which is recommended in the assessment of RA disease activity (Anderson *et al,* 2012). Nor is it included in the 2010 ACR/EULAR classification criteria (Aletaha *et al,* 2010), revised remission criteria (Felson *et al,* 2011), or the ACR core set (Felson *et al,* 1993). Therefore stiffness is not addressed in any standardised assessment of RA disease activity used in clinical or research settings. A core set is a list of measures that must be assessed in trials which standardises assessment and enhances comparison of results across studies (Felson *et al,* 1993; van Riel and van Gestel, 2000). As RA presents with a diverse range of signs and symptoms and no single 'gold standard' measure could assess all people with RA (Pincus, 2005; van Riel and van Gestel, 2000), different sets of measures were used in different trials (van Riel and van de Putte, 1994) which led to difficulties when trying to compare trial results. Therefore, a standardised core set of measures was developed for use in all clinical trials (Felson *et al,* 1993). The ACR core set was developed following discussion by a committee of experts, presentation and consensus gained at the OMERACT conference, and further subsequent discussion. It contains seven measures; three physician assessed (SJC, TJC, global assessment), three patient assessed (physical function, pain, PtG assessment), and one acute phase reactant (ESR or CRP). A radiograph is used if the study lasts longer than one year. MS was a candidate measure however, it was omitted because it was not sensitive to change (Felson *et al,* 1993). Although the RA core set does not restrict other measures such as stiffness being assessed in trials (Felson *et al,* 1993), it has likely had implications on the use of stiffness as an outcome. The lack of requirement for stiffness to be assessed in all trials may explain the decline in the reporting of stiffness in trials identified above (Labitigan *et al,* 2010) as the core set was published between the defined time-points (Felson *et al,* 1993). This may have also limited further research into understanding and assessing this symptom. Despite not being included in the ACR core set, stiffness

is a relevant outcome that is used by clinicians when making treatment decisions (Kirwan *et al,* 1984; Soubrier *et al,* 2006), and is particularly relevant in certain areas of research. For example, MS is a key outcome in research into timed-release glucocorticoid treatments (e.g. Buttgereit *et al,* 2008).

1.6 Current understanding of stiffness in RA

1.6.1 Stiffness pathophysiology

It has been well documented that RA symptoms including stiffness demonstrate highest activity in the morning. Early RA research using objective stiffness assessment (as described in Ingpen and Kendall, 1968) reported higher levels of stiffness at 6am compared to 6pm (exact values not reported, Ingpen, 1968). Similar findings have been replicated using subjective assessment in the form of MS duration (Dekkers *et al,* 2000) and MS severity (Harkness *et al,* 1982). However, in the latter study, mean MS severity (assessed on a 0 (absent) to 3 (severe) scale) was highest at 6am (mean 2.0) and lowest at 12pm (mean 0.9), although mean values at 6pm were only slightly higher (mean 1.2) (Harkness *et al,* 1982). Variation in symptoms is thought to be linked to circadian rhythms which drive biological processes such as inflammation (Straub and Cutolo, 2007; Perry *et al,* 2009). Relationships between a number of pro-inflammatory cytokines (e.g. IL-6) and RA symptoms, such as MS have also been suggested. Arvidson *et al.* (1994) demonstrated that the mean serum levels of IL-6 decreased significantly in RA patients between 7.30am (95.9 picograms/millilitre (pg/ml)) and 10.30pm (27 pg/ml, $p<0.001$). Similarly, Perry *et al.* (2009) demonstrated significantly higher mean IL-6 levels at 7.15am (35 pg/ml) than at 10pm (64 pg/ml, $p<0.001$) in an overnight study (thus not truly investigating circadian rhythms). Both studies indicated an overnight rise in IL-6 with a peak in the early morning (Arvidson *et al,* 1994; Perry *et al,* 2009) which is similar to variation in stiffness (e.g. Dekkers *et al,* 2000). Therefore, as the circadian rhythms of IL-6 and MS are similar, the circadian variation of MS may be related to increased serum IL-6 in RA (Arvidson *et al,* 1994).

1.6.2 The patient perspective of stiffness in RA

In the early stages of planning this research, little was known about the RA patient experience of stiffness. There was one study focusing solely on qualitatively understanding the patient experience of stiffness that aimed to develop a patient-centred definition (Lineker *et al,* 1999). Twenty-four people with RA took part in individual semi-structured interviews. The interview topic guide asked participants to

describe their MS in relation to how it affected their behaviour; what MS was affected by; its duration, severity, location, and variability; and its relationship to other symptoms and the previous day's activities. The topic guide was pretested with RA patients prior to use. However, it asked specifically about MS and did not appear to provide participants the opportunity to describe stiffness in their own words. Data were qualitatively analysed and formulated into statements about the characteristics of MS which were subsequently posted to the original interview participants who were able to respond to each statement on a 5-point Likert scale (strongly agree to strongly disagree). This work resulted in a definition of MS as "slowness or difficulty moving the joints when getting out of bed or after staying in one position too long, which involves both sides of the body and gets better with movement" (Lineker *et al,* 1999, p.1105). This definition is unclear as it includes stiffness 'after immobility' which may not necessarily be MS as specified. Furthermore, the term MS was retained despite patients describing stiffness as present at other times of day or throughout the day. In addition, this paper was performed almost 20 years ago, since when there have been substantial changes in RA management and treatment (Smolen *et al,* 2010) and therefore possibly changes in stiffness experience.

1.7 Stiffness treatment and management

Within the literature a number of pharmacological and non-pharmacological interventions are suggested for stiffness. However, there is no standard treatment and management approach.

1.7.1 Pharmacological treatment and management

In terms of pharmacological interventions, recent research has explored chronotherapy where treatments are coordinated with circadian rhythms (Buttgereit *et al,* 2013). Here modified-release tablets (MRT) or timed-release tablet (TRT) are used, which release the active ingredient four hours after ingestion (Buttgereit *et al,* 2013). In RA, MRT or TRT glucocorticoid treatments have been used to target symptoms of inflammation such as stiffness, before inflammatory activity starts. Observational trials have supported the overnight increase in IL-6 (Section 1.6.1) in people with RA. In a study involving nine RA patients, 24-hour blood sampling was conducted following two weeks of TRT prednisone. This demonstrated reductions in the amount of IL-6 present following TRTs (Clarke *et al,* 2011). The circadian administration of prednisone in RA (CAPRA) 1 and 2 trials, in which MS was a primary outcome measure, subsequently demonstrated reductions in MS using TRT

prednisone in large RCTs (Buttgereit *et al,* 2008; Buttgereit *et al,* 2013). In CAPRA 1, a night-time dose of timed-release glucocorticoid targeted at suppressing the early morning rise in IL-6, reduced MS duration by a mean 22.7%, compared to 0.4% in patients receiving the same dose in the morning (Buttgereit *et al,* 2008). In CAPRA 2, significantly less MS severity, duration, and daytime recurrence were reported in the TRT group compared to the placebo group (Buttgereit *et al,* 2013). These studies assessed stiffness using a diary where patients recorded the time of waking, whether stiffness was present, and the time of resolution of stiffness. The time in minutes was then calculated as the difference between the time of resolution of MS and the time of waking. Patients were also asked to record the severity of stiffness in the morning and whether stiffness recurred during the day (reported in the evening).

Early research suggestions of the effective use of glucocorticoids given in the evening (e.g. de Silva, Binder, and Hazleman, 1984) are contrary to early recommendations for daily doses being given in the morning (e.g. DiRaimondo and Forsham, 1958; Arvidson *et al,* 1997). Therefore glucocorticoids given in the morning continue to be routine as the evidence available for evening treatment is ambiguous (Kirwan, 2011). However, following the combination of advances in tablet technology and research demonstrating IL-6 and MS can be reduced by TRTs (e.g. Buttgereit *et al,* 2008; Clarke *et al,* 2011; Buttgereit *et al,* 2013), this may be changing and this rapidly developing area of research will continue using stiffness as an outcome.

1.7.2 Non-pharmacological treatment and management

There are a number of treatments that have been suggested to relieve stiffness, including exercise, hydrotherapy, splinting and heat (Hill, 2006). However, there is very little research regarding the effectiveness of these. In a systematic review looking at the effects of compression gloves in arthritis, four RCTs were identified (Hammond, Jones and Prior, 2015). Two reported significant reductions in self-reported stiffness, however this was also reported in the placebo glove. Furthermore, the studies were identified as being poor quality and the review concluded that there is inconclusive evidence for the use of compression gloves in RA (Hammond, Jones and Prior, 2015). Another review of therapy gloves identified eight articles, seven of which reported on stiffness (Nasir, Troynikov and Massy-Westropp, 2014). Of these, six reported an improvement in MS duration or severity following the use of therapy gloves. However, the studies were generally of poor quality with little description of the therapy glove used (Nasir, Troynikov and Massy-Westropp, 2014). It was recognised in the one paper where no improvement was found, that this could be a

result of difficulty in stiffness assessment as patients were unable to differentiate between ‘mild’ or ‘moderate’ stiffness on the measurement scale used (Swezey *et al,* 1979). This suggests that difficulty in the area of stiffness measurement may contribute to the limited research in this area.

1.8 Stiffness in other rheumatic diseases

Stiffness is not only relevant in the context of RA, it is also recognised by patients and clinicians in a number of other rheumatic diseases including polymyalgia rheumatica, AS, and OA. One of the criticisms of the 1987 classification criteria that includes MS (Arnett *et al,* 1988) was that it was not specific enough and can lead to false positive classification of RA in patients with other inflammatory conditions (Levin *et al,* 1996). The symptoms of pain and stiffness in RA, PMR, and AS appear to be highest in the early morning (Spies *et al,* 2010) which may indicate a shared pathology. However, there is little work exploring similarities and differences in the patient experience of stiffness in these different conditions.

1.8.1 Polymyalgia rheumatica (PMR)

PMR is a condition characterised by pain and stiffness, particularly of the shoulders and hips (Arthritis Research Campaign, 2005). MS >45 minutes is heavily weighted in the scoring algorithm in the PMR ACR/EULAR classification criteria (Dasgupta *et al,* 2012) and it features in the outer core of the provisional core domain set for PMR (Helliwell *et al,* 2016). Patients with PMR endorse stiffness as an important treatment outcome (Mackie *et al,* 2014) and a recent qualitative study developed understanding of the meaning, experience and impact of stiffness in PMR patients (Hughes *et al,* 2012). In this work, MS or EMS were not typically described by patients, who instead discussed stiffness affecting them 24 hours a day with worsening at night and in the early morning. Patients also described a close relationship between stiffness and pain (Hughes *et al,* 2012). This is an important step towards a clearer understanding of stiffness in PMR.

1.8.2 Ankylosing spondylitis (AS)

AS is the most common of the spondyloarthritides and is an inflammatory condition that affects the spine, causing pain and stiffness (Hill, 2006). Stiffness is a feature in the modified criteria for AS (van der Linden, Valkenburg, and Cats, 1984) and is included in two items within the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) (Garrett *et al,* 1994), which is the gold standard AS disease activity

assessment. Qualitative work into the patient experience of AS found that along with pain and fatigue, stiffness was discussed by participants and featured in all three identified themes which captured different life situations across levels of symptom management and requiring behaviour adjustment: ordinary life, slowed-down life and disrupted life (Mengshoel, 2008). However, the study involved 12 AS patients, eight of whom were women, which does not reflect an AS population that has a male to female ratio of 3:1 (Hakim, Clunie and Haq, 2011). There appears to be no current work specifically exploring stiffness in AS.

1.8.3 Osteoarthritis (OA)

OA is the most common form of rheumatic disease in the UK causing destruction of the hyaline cartilage of bone surfaces and overgrowth of bone, resulting in joint pain and stiffness in affected joints (Arthur and Hill, 2006). Stiffness is assessed in a number of composite OA assessments such as the Western Ontario McMaster Osteoarthritis Index (WOMAC) (Bellamy *et al,* 1988) and the Australian/Canadian Osteoarthritis Hand Index (AUSCAN) (Bellamy *et al,* 2002). There appears to be no current work specifically exploring the patient experience of stiffness in OA, but there is qualitative work exploring the patient perspective in relation to symptoms more broadly. One study investigated how patients manage symptoms of pain, aching or stiffness using focus groups (MacKay *et al,* 2014a) and interviews (MacKay *et al,* 2014b), in total involving 51 people with self-reported OA or knee symptoms. Three core themes were identified following the focus groups including control of symptoms, seeking solutions, and active management, and two core themes using the combined dataset, including disrupted physical, emotional and social life and altered way of thinking about the body and self. This work highlights the engagement of patients in disease management strategies but also the broad impact, in the form of disruption that symptoms had. Another study involving the completion of self-reported questionnaires by 80 patients with clinically diagnosed OA, explored patient expectations about future symptoms (Dwek *et al,* 2015). Results indicated that generally patients were positive about future symptoms and expected the same or less pain and stiffness. Patients who predicted stiffness in one year to be better than at present also reported higher levels of behavioural engagement which reinforces earlier work (MacKay *et al,* 2014a; 2014b) regarding engagement in symptom control and management.

1.9 Stiffness in other populations

Stiffness is also relevant in other populations, unrelated to rheumatology.

1.9.1 Older populations

Stiffness has been reported as relevant for older individuals within the general population. One study compared symptoms in individuals aged >50 years in the general population, with symptoms in people with RA (Sokka *et al,* 2007) to identify the proportion in each population that met the ACR remission criteria (Pinals *et al,* 1981) or the OMERACT criteria for minimal disease activity (Wells *et al,* 2005). The study was based on a self-reported questionnaire to which 3105 people responded (80% response rate), of whom 1705 were RA patients, and 1400 were aged >50 years and in the general population. Over 15 minutes of MS was reported by 64.7% of RA patients and 36.6% of those aged >50 years in the general population (Sokka *et al,* 2007) indicating that MS is relevant in the general population, although not to the same extent as in an RA population.

1.9.2 Stiff person syndrome

Stiff person syndrome is a rare but debilitating condition that causes stiffness and spasms, and results in a loss of independence and increased risk of falls (Dalakas, 2009). Although the underlying process of stiff person syndrome is unclear, it is thought to be related to the autoimmune system, in particular the glutamic acid decarboxylase autoantibody which is linked to diabetes mellitus, which is often present in patients with stiff person syndrome (Hadavi *et al,* 2011). Although the symptom of stiffness is consistent with RA, the underlying mechanisms, although poorly understood in both conditions, appear to have different drivers.

1.9.3 Phantom limb syndrome

Stiffness has also been described by individuals with amputations who experience phantom limb. In a small study involving three people with RA and lower limb amputation, it was reported that patient-reports of stiffness were similar in their amputated and non-amputated limbs. As a result it was suggested that changes in the central nervous system may be the driving mechanism behind subjective patient-reports of stiffness (Haigh *et al,* 2003). However, it has been argued that this proposal does not explain the evidence regarding the objective assessment of stiffness (Helliwell, 2004).

Chapter 2: Stiffness assessment and PROM development

The previous chapter provided an introduction to RA and how it is managed and assessed. This chapter will address stiffness assessment, the purpose and use of PROMs and the literature surrounding their development. It will identify currently available stiffness PROMs, and their measurement property evidence.

2.1 Stiffness assessment in RA

The traditional biomedical model of healthcare focused on physical healing of biological dysfunction, therefore traditional measurement concepts focused on objective measurement (Elasy and Gaddy, 1998). Early work into stiffness assessment used engineering principals to design apparatus called arthrographs to measure the elasticity, inertia and plasticity of joints. Scott (1960) used a spring applied to the end of a finger and assessed the distance that the finger moved from a horizontal position. Ingpen and Kendal (1968) used a lever based mechanism and assessed the time taken for the finger to move the required distance. Wright and Johns (1960, 1961) used apparatus based on a pulley and lever system and assessed the force required to move the finger of individuals with and without connective tissue disease. This work concluded that these methods allowed the measurement of physical joint stiffness. However, no difference was demonstrated between participants with and without connective tissue disease (Wright and Johns, 1961). The use of such early apparatus was restricted in its practical application because it was limited to use in the hands only, and it was bulky and uncomfortable for patients (Wright, Dowson, and Longfield, 1969). However, the development of lighter and more compact apparatus could be used in applied settings (Howe, Thompson, and Wright, 1985). Success with such apparatus was variable, as shown in one study which demonstrated that stiffness was no greater in RA patients than healthy controls (Helliwell, Howe and Wright, 1988). Despite this, further work provided a better understanding about the influence of muscle wasting and the positioning of the joint during measurement (Helliwell, Howe and Wright, 1987a), which led to a new measurement process that demonstrated increased stiffness in RA patients compared to healthy controls (Helliwell, Smeathers, and Wright, 1994). Although a positive conclusion was reached about the effectiveness of measures of physical stiffness, it has been suggested that this was outweighed by the effort of collecting this information in comparison to other tests of inflammation such as CRP (Helliwell *et al,* 2007). Furthermore, much of the literature regarding the assessment of physical stiffness is now rather dated which questions the appropriateness of its

application to todays population. Although more recently, the issue of physical stiffness measurement has been revisited by research focused on the design of a 'data glove' to quantify joint stiffness and range of movement of the hand by measuring finger joint kinematics (Connolly *et al,* 2012). Although this does not solve the problem of stiffness affecting more than just the hands, it could provide a solution for the practical application of such a device in clinical, community or home based settings. Another recent abstract has reported using electromyography, a way of measuring electrical activity in muscles, to objectively assess stiffness (Mengi *et al,* 2014).

Another method of obtaining symptom assessment information is by asking patients directly. MS duration was included in early disease activity assessment (e.g. Lansbury index, Lansbury, 1958) and classification criteria (Arnett *et al,* 1988), and MS severity was used in research contexts (e.g. Harkness *et al,* 1982). Patient-reported assessment reflects the increased importance of the use of PROs (de Wit *et al,* 2011) and has clear practical advantages over the bulky, apparatus based strategies described above. Another important justification for patient-reported assessment of stiffness is the lack of relationship between physical stiffness measures and patient-reported methods. A study comparing different stiffness assessment methods including a physical stiffness arthrograph and patient-reported assessment of MS duration and severity showed poor correlations and the authors concluded that the objective assessment methods did not relate to the subjective patient reports (Rhind, 1988). These finding were replicated in other work suggesting that physical stiffness is different to patient-reported stiffness (Helliwell, Howe and Wright, 1987b; Helliwell, Howe and Wright, 1988).

2.2 Patient reported outcome measures (PROMs)

2.2.1 What are PROMs?

PROMs are an assessment of a patient's health condition that comes directly from a patient, without any interpretation from another individual (US Department of Health and Human Services, Food and Drug Administration (USDHHS FDA), 2009). PRO is a term used to describe the concept of interest (e.g. disability) while PROM refers to a specific questionnaire that represents that PRO concept (e.g. HAQ, Fries *et al,* 1980) (Patrick *et al,* 2011a). PROMs are particularly relevant when measuring concepts that are best understood by patients (USDHHS FDA, 2009), such as disease symptoms. They are vital in chronic conditions, where the evaluation of

therapies or interventions is more relevant than survival (Patrick *et al,* 2007). Here, some effects of therapies or interventions may only be known to the patient and this information would not be captured if not assessed directly (USDHHS FDA, 2009). PROMs also reflect aspects relating to patient priorities for therapies or interventions more so than captured in other clinician assessed outcomes (Leidy and Vernon, 2008).

2.2.2 Use of PROMs

The recent focus on the use of PROMs highlights the increasing emphasis on incorporating patients and the public in all aspects of healthcare (Department of Health, 2010a) and is representative of the ideological shift away from patients being passive receivers of a service, towards patients as empowered and active participants in all aspects of care (Foot *et al,* 2014). As PROMs capture the patient perspective, they are particularly important in facilitating the involvement and engagement of patients in healthcare situations such as decision making (Frost *et al,* 2007). They are also valuable for highlighting what is relevant to them as a patient, including patient-relevant symptoms (Kirwan *et al,* 2007). A specific example of the importance of patient involvement is the involvement of RA patients at OMERACT meetings which led to recognition of fatigue as an important patient symptom which is now measured alongside the core set in all RA studies as recommended by international consensus (Kirwan *et al,* 2007). As PROMs are only effective if they genuinely capture the patient perspective (Kerr, Nixon and Wild, 2010), it is vital that patients are involved in their development, specifically in relation to enhancing content validity (USDHHS FDA, 2009). Despite this, the development of content validity is often neglected (Patrick *et al,* 2007), and some PROMs are developed with little or no input from patients (Section 2.4) and therefore do not include the experiences and perspective of the user that is essential for content validity.

2.3 PROM development theory

2.3.1 Concepts of measurement

For a measurement tool to be useful it must demonstrate appropriate measurement properties including validity, reliability, internal consistency, ability to detect change, floor and ceiling effects and interpretability (Terwee *et al,* 2007). Measurement properties and their assessments are described below.

2.3.1.1 Validity

Broadly, validity is the degree to which a questionnaire measures what it intends to measure (Scientific Advisory Committee of the Medical Outcomes Trust (SACMOT), 2002; Frost *et al,* 2007). Face and content validity refer to whether a measure looks appropriate (Streiner and Norman, 2008). Face validity refers to whether users perceive the instrument to capture the relevant information, while content validity looks at whether the instrument captures the appropriate and full range of relevant content (Frost *et al,* 2007; Streiner and Norman, 2008). The development of face and content validity requires qualitative methods with the relevant populations to develop and select appropriate items (Frost *et al,* 2007; Terwee *et al,* 2007; USDHHS FDA, 2009). Criterion validity refers to how the questionnaire relates to other valid measures that assess the same concept (Streiner and Norman, 2008) or to a known 'gold standard' (USDHHS FDA, 2009). This is difficult to assess as there is rarely a 'gold standard' available for comparison (Frost *et al,* 2007), as the lack of other valid measures is often the reason for the development of a new scale. Construct validity is the extent to which the measure relates to theoretically relevant constructs (Frost *et al,* 2007; Terwee *et al,* 2007). Construct validity can target different aspects including convergent (demonstration of relationships where expected), divergent (demonstration of no relationship where expected), and discriminant (ability to distinguish between expected or known groups) validity (Streiner and Norman, 2008). A common approach to examine construct validity is to compare responses between the instrument to be tested and responses to measures that capture theoretically related concepts (SACMOT, 2002). Each aspect of construct validity should be performed using correlations to test specific and predefined hypotheses (Terwee *et al,* 2007).

2.3.1.1.1 Definition of correlation cut-offs

To assess aspects of validity identified in Section 2.3.1.1, Pearsons (parametric data) or Spearman's rank order (non-parametric data) correlations can be used. A correlation of 1 or -1 indicates a perfect correlation while 0 indicates no correlation (Pallant, 2010). However, there are different recommendations regarding the strength of correlations required to define levels of acceptable validity (Table 2.1). For the purposes of validity testing in this study, correlations of r=<0.5 were defined as weak, r=0.5-<0.7 as moderate and r=≥0.7 as strong. Broadly, strong correlations would be expected between measures assessing the same construct and weak correlations between measures addressing different constructs (Frost *et al,* 2007).

Table 2.1: Different correlation cut-off definitions

Authors	Correlation cut-off definition
Cohen (1988, cited in Pallant, 2010)	Small: *r*=0.10-0.29 Medium: *r*=0.30-0.49 Large: *r*=≥0.50
Pett, Lackey and Sullivan (2003)	Weak: *r*=≤0.29 Low: *r*=0.30-0.49 Medium: *r*=0.50-0.69 Strong: *r*=0.70-0.89 Very strong: *r*=≥0.90
Dancey and Reidy (2007)	Weak: *r*=0.1-0.3 Moderate: *r*=0.4-0.7 Strong: *r*=0.7-0.9

2.3.1.2 Reliability

Reliability is the degree to which a measure produces the same score each time it is administered when the measurement construct has not changed (SACMOT, 2002; Frost *et al,* 2007). Reliability can be assessed using test-retest in a population that has not changed (i.e. is the measure reproducible?) (Field, 2009).

2.3.1.3 Internal consistency

Internal consistency is an aspect of reliability concerned with the homogeneity of the items that make up an instrument (Field, 2009; DeVellis, 2012). Internal consistency looks at the relationships between items on the basis that items measuring the same concept should be consistent with each other. The advantage of internal consistency is that unlike test-retest, it can be generated from a single administration of the tool (Streiner and Norman, 2008). Internal consistency can be assessed using Cronbach's alpha coefficient (DeVellis, 2012) where values of >0.7 are acceptable but values >0.8 are preferred (Pallant, 2010).

2.3.1.4 Ability to detect change

Ability to detect change or responsiveness is concerned with a tool's ability to identify change (Terwee *et al,* 2007) for example, when the patient experience changes in the attribute being assessed, does the instrument score reflect that change? Definition of the minimally important clinical change expected and receiver operator curves (ROC) can explore this (Terwee *et al,* 2007).

2.3.1.5 Other concepts of measurement

Other concepts of measurement include floor and ceiling effects and interpretability. Floor and ceiling effects enable exploration of whether many participants respond to items using the highest or lowest categories and identify if more response options are required (Terwee *et al,* 2007; USDHHS FDA, 2009). Interpretability focuses on whether the quantitative scores from the measure relate to appropriate qualitative meaning (Terwee *et al,* 2007).

2.3.2 Meeting the OMERACT filter

The OMERACT Filter is a method of ensuring quality in PROMs for use in research and clinical settings. It is based around three concepts; truth, discrimination and feasibility (Boers *et al,* 1998). Truth relates to validity and whether the measure is unbiased and measures what it intends to measure. Discrimination relates to reliability and sensitivity and whether the measure can discriminate (e.g. patients with active or inactive disease). Feasibility relates to whether the measure is useful in applied situations (e.g. understandable and time efficient). The Filter relates to all parties involved in assessment including patients and depending on the purpose of the tool, researchers and/or clinicians (Boers *et al,* 1998). The original Filter has recently been recently updated. Filter 2.0 maintains the emphasis on the three original concepts but puts these in the context of updated philosophical and methodological approaches to health assessment (Boers *et al,* 2014).

2.3.3 PROM development guidelines

The FDA have produced guidelines on the evaluation, modification and development of PROMs, specifically in the context to support labelling claims (USDHHS FDA, 2009). Although the focus on supporting labelling claims is not specifically relevant in this research, the use of a rigorous framework for item development is important. The development of content validity is also vital and this has been the focus of the International Society for Pharmaceconomics and Outcomes Research (ISPOR) PRO Content Validity Good Research Practices Task Force guidelines (Patrick *et al,* 2011a, Patrick *et al,* 2011b). The use of such guidelines enable a rigorous and systematic process of PROM development.

2.3.4 Test theory

There are different approaches to PROM development including classical test theory (CTT) and item response theory (IRT). CTT is the traditional approach to outcome

measure development (Streiner and Norman, 2008) and is grounded in the idea that a participant's observed score is the result of true score plus error (DeVellis, 2012). CTT is a broadly accepted and well used in scale development (Streiner and Norman, 2008). However, there are limitations with CTT such as its dependency on the sample involved in development, which means that to apply items to a different sample requires re-testing it in each different sample. IRT includes models such as Rasch (Streiner and Norman, 2008), and was developed to overcome the limitations with CTT. For example, unlike in CTT, IRT models are not specific to the development population therefore the advantage of IRT is its invariance property (Streiner and Norman, 2008). However, IRT has its own associated limitations including that its invariance properties are not always demonstrated and differences between populations have been reported (Cook, Eignor and Taft, 1988; Miller and Linn, 1988).

2.4 Current stiffness PROMs

A systematic literature review investigating current stiffness assessment was recently performed (van Tuyl, Lems and Boers, 2014), aiming to identify currently available stiffness assessment measures and summarise their measurement properties. However, its focus was narrow, specifically looking at the measures available to assess RA patients in low disease activity or remission states. In order to identify currently available stiffness PROMs used across all RA disease activity states this systematic review was updated and its focus expanded (Section 2.4.1).

2.4.1 Systematic literature review of stiffness PROMs

2.4.1.1 Objectives

This systematic literature review aimed to identify currently available stiffness assessment measures and summarise the evidence of their measurement properties.

2.4.1.2 Methods

To retain consistency, methods were based on the systematic review performed by van Tuyl, Lems and Boers (2014).

2.4.1.2.1 Search strategy

The search strategy included three key searches concerning: 1) construct (stiffness); 2) population (adults with RA); 3) instrument (PROM). These topics were searched with a validated sensitivity and exclusion filter, designed to identify studies on measurement properties of instruments (Terwee *et al,* 2009). As the search strategy

was provided it was possible to replicate and update the search in accordance with the original review (van Tuyl, Lems and Boers, 2014), performed in PubMed on 20/11/2012. The updated search was also performed in PubMed specifying dates from 20/11/2012 to 20/11/2013.

2.4.1.2.2 Selection criteria and data collection

Four rounds of review were performed to identify articles. Round one involved screening titles and abstracts to identify papers that included reports of the appropriate construct (stiffness) and PRO measurement, development or validation. This was performed by one researcher (Halls). Round two involved screening those titles and abstracts to identify papers that included reports of the appropriate population (adults with RA) and stiffness measurement properties. This was performed independently by Halls and a member of the supervisory team (SH). Any disagreements were discussed to guide decision making. Round three involved screening those full articles (Halls). Papers were selected for further consideration if they reported on stiffness in one of three ways: 1) stiffness as an outcome in relation to other core set disease activity measures; 2) the development of a stiffness PROM; 3) a comparison of two or more different tools to measure aspects of stiffness or between a questionnaire that includes a stiffness item against another stiffness item. The reference lists of articles included in round three were also reviewed. Round four involved identifying the available stiffness PROMs and extracting information regarding their measurement properties, performed by Halls and discussed with a member of the supervisory team (SH). The data extraction form (Appendix A) was developed based on an example used in a study with similar aims (Hewlett, Hehir and Kirwan, 2007), and quality criteria for evaluating questionnaires (Terwee *et al,* 2007).

2.4.1.3 Results

2.4.1.3.1 Articles

Nineteen full articles were included in the final selection round of the review. Three articles were found in the updated search and the other 16 were taken from the original review (van Tuyl, Lems and Boers, 2014) (Figure 2.1). During round two, the reviewers had different opinions about six papers (Shirinsky *et al,* 2013; Zakeri *et al,* 2013; Jastrząbek *et al,* 2013; Cutolo *et al,* 2013; Wiesinger *et al,* 2013; Buttgereit *et al,* 2013). Given the uncertainty, these were retained for full review in round three. Although the reference lists were reviewed and two papers of interest were identified

(Lineker *et al,* 1999; Buttgereit *et al,* 2008), they did not contain relevant information but have been discussed elsewhere (Sections 1.5.4 and 1.7.1).

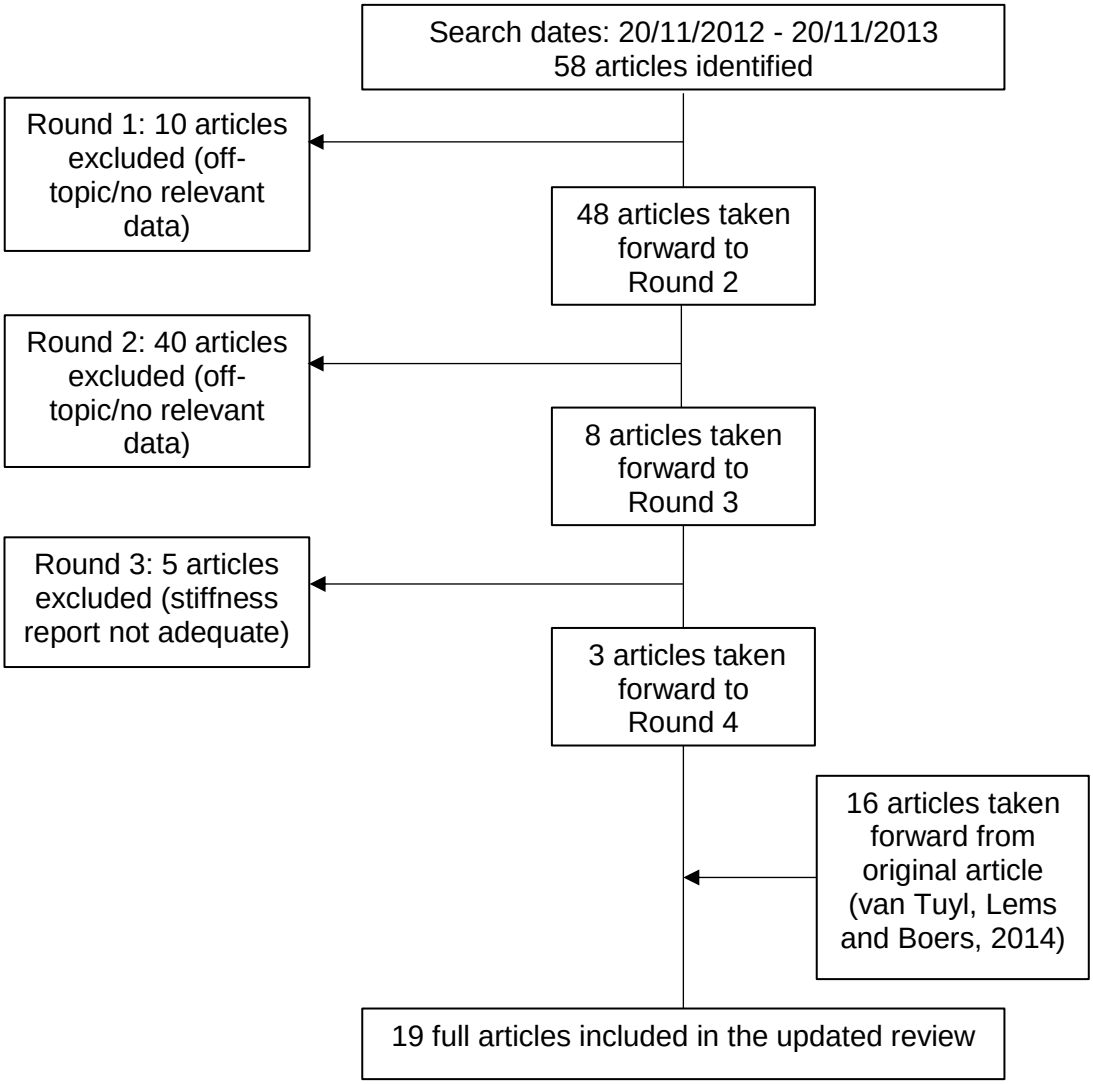


Figure 2.1: Flow diagram showing article selection process

2.4.1.3.2 Stiffness PROMs

The stiffness assessment measures identified in the 19 articles (Table 2.2) all assessed stiffness from the concepts of duration or severity. Two studies did not define the concept that was assessed (Borstlap *et al,* 1995; Wiesinger *et al,* 2013). Stiffness assessment measures predominantly assessed MS or EMS. There were two exceptions; one investigated the severity of stiffness after sitting, lying down or resting during the day with clearly defined wording (Wolfe, 1999). Another assessed 'starting stiffness after a time of rest' which appeared to be a severity item given by the anchors but the item wording was not defined (Leeb *et al,* 2003). Despite the few concepts assessed, there was considerable variation not only in respect to the

wording of the question and response options, but also in relation to format (e.g. VAS or NRS) and timeframe (e.g. today, last 48 hours, last week). Many articles did not define the wording or other aspects of the items used. As a result it is not clear how stiffness questions were asked or reported. This means that not only are there many different versions of questions but where different questions are used, the exact format in which they are used is unknown, which further limits comparison across studies or replication.

Table 2.2: Description of current stiffness assessment measures

Authors	Item concept	Stem question	Response options/anchors	Timeframe	Patient sample
Rhind, Unsworth and Haslock, 1987	Severity of MS	Exact wording unclear	10cm VAS: No stiffness to Very severe stiffness	At the time of interview and on the day of interview	95 RA
	Severity of MS	Exact wording unclear	11-point NRS: 0 (No stiffness) to 10 (Very severe stiffness)		
	Severity of MS	Exact wording unclear	5-point verbal scale: No stiffness, mild, moderate, severe, very severe stiffness		
	Duration of MS	How long did it take for your stiffness to begin to ease after you got out of bed this morning?	Minutes	Today	
Hazes, Hayton, and Silman, 1993	Severity of MS	Exact wording unclear	10cm VAS: No stiffness to Very severe stiffness	Today	93 RA+46 NIC
	Severity of MS	Exact wording unclear	11-point NRS: 0 (No stiffness) to 10 (Very severe stiffness)	Today	
	Duration of MS	How long does you MS last until it begins to improve?	Minutes	Today	
	Duration of MS	How long does your MS last until maximum improvement occurs?	Minutes	Today	
	Duration of MS	How long does it take you to get going properly?	Minutes	Today	

Authors	Item concept	Stem question	Response options/anchors	Timeframe	Patient sample
Hazes *et al,* 1994	Duration of MS	Waking to first improvement of MS	Reported in minutes and categorised into: >1 hour; 1-3 hours; ≥3 hours	Today	49 RA
	Duration of MS	Waking to maximum improvement of MS		Today	
	Duration of MS	Waking to complete disappearance of MS		Today	
	Duration of MS	Getting up to maximum improvement of MS		Today	
	Duration of MS	Getting up to first improvement of MS		Today	
	Duration of MS	Getting up to complete disappearance of MS		Today	
Ward, 1994	Duration of MS	Patients were asked to report if they experienced morning stiffness and if so to estimate how long it typically lasted. Exact wording unclear	Minutes. Exact options unclear	Unclear	24 RA
Buchbinder *et al,* 1995	Duration of MS	Patients asked to record time of awakening, time of arising, and time of cessation of MS. Exact wording unclear	Minutes (duration measured as time from awakening)	The day preceding the clinic visit	142 RA
Borstlap *et al,* 1995	No mention of duration or severity	Exact wording unclear	10cm VAS: the lower the score, the more favourable the patients' condition. Anchor wording unclear	Unclear	62 OA+35 RA

Authors	Item concept	Stem question	Response options/anchors	Timeframe	Patient sample
Vliet Vlieland *et al,* 1997	Duration of MS	How long does your morning stiffness last from waking until maximum improvement occurs?	Minutes (cut-off at 240 minutes)	Today	63+39 trial RA
	Severity of MS	Exact wording unclear	10cm VAS: None to Very severe	Today	
Houssien, McKenna and Scott, 1997	Duration of EMS	Exact wording unclear	Minutes	Unclear	200 RA
Wolfe, 1999	Severity of MS	How severe has your stiffness been after you first woke up in the morning?	Both validated in 5-point Likert scale (0=none, 1=mild, 2=moderate, 3=severe, 4=extreme), 100mm VAS (0 (none)-10 (extreme)) and 11-point NRS (0 (none)-10 (extreme)). Item wording begins with: Think about stiffness (not pain) you felt during the last 48 hours caused by the arthritis in your knee to be injected. Stiffness is the sensation of decreased ease in moving your joint (Bellamy *et al,* 1988)	Last 48 hours	1013 RA, 625 OA, 531 FM
	Severity after sitting, lying down or resting during the day	How severe has your stiffness been after sitting or lying down or while resting later in the day?			
Fransen *et al,* 2000	Duration of MS	Were your joints stiff when you woke up today? No/Yes If yes, how long did this extra stiffness last?	7-point Likert scale (0=none, 1=<30 minutes, 2=30 minutes to an hour, 3=1-2 hours, 4=2-4 hours, 5=more than 4 hours but less than all day, 6=all day)	Today	584 RA
Sarzi-Puttini *et al,* 2002	Duration of MS	Exact wording unclear	Minutes reported on a 100 mm VAS? Anchor wording unclear	Unclear	105 RA

Authors	Item concept	Stem question	Response options/anchors	Timeframe	Patient sample
Leeb *et al,* 2003	Daily MS severity	Exact wording unclear	100mm anchored VAS: no stiffness to unbearable stiffness	Unclear	103 RA+69 OA
	Starting stiffness after a time of rest	Exact wording unclear	100mm anchored VAS: no stiffness to unbearable stiffness	Unclear	
	Duration of MS	Exact wording unclear	Minutes	Unclear	
Yazici *et al,* 2004	Duration of MS	Exact wording unclear	Reported in minutes and subsequently categorised into four groups: 0, 1-15, 16-59, and 60+	Unclear	337 RA
Westhoff *et al,* 2008	Severity of MS	Exact wording unclear	11-point NRS: No morning stiffness at all to Extremely severe morning stiffness	Unclear	916 RA
	Duration of MS	Exact wording unclear	Minutes	Unclear	
Khan *et al,* 2009	Duration of MS	From time of waking to time of max improvement	Reported in minutes and subsequently categorised into: None (0 minutes), mild (1-30 minutes), moderate (31-60 minutes), severe (>60 min)	Last week	5439 RA
El Miedany *et al,* 2010	Duration of MS	'Over the last week when you awakened in the morning, did you feel stiff? Please indicate the number of minutes, or hours until you are as limber as you will be for the day.'	Minutes/hours	Not defined for question	264 RA, 123 PsA+75 IBD

Authors	Item concept	Stem question	Response options/anchors	Timeframe	Patient sample
Wiesinger *et al,* 2013	No mention of duration or severity	Exact wording unclear	Unclear	Unclear	451 RA
Jastrzabek *et al,* 2013	Duration of MS	Exact wording unclear	Minutes	Unclear	40 RA
Lie *et al,* 2014	Severity of MS	How would you describe the overall level of morning stiffness you have had from the time you wake up? (Q5 from BASDAI; Garrett *et al,* 1994)	10cm VAS: 0 (None) to 10 (Very Severe)	Unclear (past week used in BASDAI)	1195 RA (stiffness data for 39% of patients)
	Duration of MS	How long does your morning stiffness last from the time you wake up? (Q6 from BASDAI; Garrett *et al,* 1994)	10cm VAS: 0 (0 hours) to 10 (2 or more hours) with marked intervals at ½ hour, 1 hour, and 1½ hours		

NIC=non-inflammatory complaint; PsA=psoriatic arthritis; IBD=inflammatory bowel disease

2.4.1.3.3 Measurement properties of stiffness PROMs

2.4.1.3.3.1 Validity

Most measurement property evidence related to validity including face, content, criterion and construct validity (Section 2.3.1.1).

2.4.1.3.3.1.1 Face and content validity

No studies reported directly on the face validity of the stiffness items. However, two studies reported on content validity (Leeb *et al,* 2003; Lie *et al,* 2014). One described the process of item generation which involved the study authors proposing relevant items which were ranked and reduced using a Delphi approach and discussions with other health professionals (Leeb *et al,* 2003). This provides content validity evidence given item development involved clinical experts. However, no patient involvement was reported. The study by Lie *et al.* (2014) was interested in the content validity of the proposed flare domains rather than stiffness specifically. Therefore studies provided limited evidence regarding the content validity of the identified stiffness items.

2.4.1.3.3.1.2 Criterion validity

Seven studies reported the relationship between different stiffness items (Rhind, Unsworth and Haslock, 1987; Hazes, Hayton, and Silman, 1993; Hazes *et al,* 1994; Vliet Vlieland *et al,* 1997; Leeb *et al,* 2003; Westhoff *et al,* 2008; Lie *et al,* 2014). These included comparisons between items assessing different concepts (duration and severity) and between items assessing the same concept (e.g. duration using different wording or timeframes). When looking at comparisons between items assessing different concepts, weak to strong correlations were reported (Table 2.3). This suggests that despite severity and duration being used interchangeably, these concepts may capture different information. However, this is difficult to tell given the variability across items.

Table 2.3: Correlations between items assessing different concepts

Authors	Items compared	Results
Westhoff *et al,* 2008	Severity of MS vs duration of MS (baseline) Severity of MS vs duration of MS (three year follow-up)	r_s=0.75, p<0.001 r_s=0.81, p<0.001
Lie *et al,* 2014	Severity of MS vs duration of MS	r_s=0.63*
Leeb *et al,* 2003	Daily MS severity vs duration of MS	r_s=0.66, p<0.0001
Vliet Vlieland *et al,* 1997	Severity of MS vs duration of MS (Study 1) Severity of MS vs duration of MS (Study 2)	r_s=0.63, p<0.001 r_s=0.62, p<0.001
Rhind, Unsworth and Haslock, 1987	Severity of MS vs duration of MS (VS) Severity of MS vs duration of MS (NRS) Severity of MS vs duration of MS (VAS)	r_p=0.42* r_p=0.41* r_p=0.46*
Hazes, Hayton, and Silman, 1993	Severity of MS vs duration of MS	Reported as 'poor'+

VS=verbal scale; *=p not reported; +=no values provided

The evidence also varied when reviewing comparisons between items assessing the same concept using different wording or timeframes. Although items assessing the same concept using different formats were highly correlated, items assessing the same concept using different timeframes correlated weakly (Table 2.4). This questions the value of using different stiffness time cut-offs and which is the most appropriate. The differences in results when excluding patients who reported stiffness all day also suggests that MS may not reflect the full patient experience of stiffness.

Table 2.4: Correlations between items assessing the same concept

Authors	Items compared	Results
Rhind, Unsworth and Haslock, 1987	MS severity NRS vs MS severity VAS	r_p=0.84*
	MS severity verbal scale vs MS severity NRS	r_p=0.81*
	MS severity verbal scale vs MS severity VAS	r_p=0.84*
	Present stiffness VS vs present stiffness NRS	r_p=0.90*
	Present stiffness VS vs present stiffness VAS	r_p=0.87*
	Present stiffness NRS vs present stiffness VAS	r_p=0.91*
	MS severity VS vs present stiffness VS	r_p=0.47*
	MS severity NRS vs present stiffness NRS	r_p=0.42*
	MS severity VAS vs present stiffness VAS	r_p=0.48*
Hazes, Hayton, and Silman, 1993	MS duration time to initial improvement vs MS duration time to maximum improvement	r=0.41*
Hazes *et al,* 1994	MS duration (time until first improvement) diary vs MS duration (time until first improvement) interview	r_p=0.68* (r_p=0.50)*^
	MS duration (time until maximum improvement) diary vs MS duration (time until maximum improvement) interview	r_p=0.42* (r_p=0.66)*^
	MS duration (time until MS disappears) diary vs MS duration (time until MS disappears) interview	r_p=-0.06* (r_p=0.88)*^

VS=verbal scale; *=p not reported; ^=excluding patients reporting stiffness all day

2.4.1.3.3.1.3 Construct validity

As construct validity explores relationships between items and theoretically related concepts (SACMOT, 2002), stiffness items would be expected to demonstrate relationships with other measures of RA including disease activity and other symptoms. It would also be expected that items could discriminate between known groups. Firstly, given that stiffness is considered an indicator of inflammatory activity in RA (e.g. Lansbury, 1956; Scott, 1986; Hazes *et al,* 1994; Soubrier *et al,* 2006) it would be expected that stiffness items would demonstrate moderate relationships with other variables reflecting inflammatory activity such as disease activity. Only two studies investigated the relationship between stiffness items and composite RA disease activity assessment (DAS28) (Westhoff *et al,* 2008; Khan *et al,* 2009). These reported weak or moderate correlations, with MS severity demonstrating marginally stronger correlations. Studies have also reported on relationships between stiffness and other disease activity measures including core set variables (e.g. SJC), again reporting weak or moderate correlations (Table 2.5).

Table 2.5: Correlations between stiffness items and disease activity

Authors	Stiffness item	Disease activity comparator	Results
Westhoff *et al,* 2008	Severity of MS (baseline)	DAS28	r_s=0.47, p<0.001
		TJC	r_s=0.31, p<0.001
		SJC	r_s=0.28, p<0.001
		CRP	r_s=0.20, p<0.001
		ESR	r_s=0.20, p<0.001
	Severity of MS (three year follow-up)	DAS28	r_s=0.58, p<0.001
		TJC	r_s=0.48, p<0.001
		SJC	r_s=0.35, p<0.001
		CRP	r_s=0.22, p<0.001
		ESR	r_s=0.21, p<0.001
Khan *et al,* 2009	Duration of MS	DAS28	r_s=0.46, p<0.001
		TJC	r_s=0.39, p<0.001
		SJC	r_s=0.33, p<0.001
		ESR	r_s=0.23, p<0.001
		PGA	r_s=0.39, p<0.001
Ward *et al,* 1994	Duration of MS	TJC	r_p=0.54, p<0.0001
		SJC	r_p=0.38, p<0.05
		ESR	r_p=0.21, p<0.05
		PGA	r_p=0.45, p<0.0001
Fransen *et al,* 2000	Duration of MS	TJC	r_s=0.36, p<0.0001
		SJC	r_s=0.25, p<0.0001
		ESR	r_s=0.17, p<0.0001
		PGA	r_s=0.36, p<0.0001
Yazici *et al,* 2004	Duration of MS	TJC	OR 1.23, 95% CI 1.14-1.30
		SJC	OR 1.05, 95% CI 1.01-1.10
		ESR	OR 1.01, 95% CI .99-1.02
Miedany *et al,* 2010	Duration of MS	TJC	r_p=0.45, p<0.01
Vliet Vlieland *et al,* 1997	Duration of MS	SJC	r_s=0.12 (Study 1), r_s=0.07 (Study 2)**
		CRP	r_s=0.10 (Study 1), r_s=0.08 (Study 2)**
		PGA	r_s=0.20 (Study 1), r_s=0.06 (Study 2)**
	Severity of MS	SJC	r_s=0.00 (Study 1), r_s=0.11 (Study 2)**
		CRP	r_s=0.01 (Study 1), r_s=0.14 (Study 2)**
		PGA	r_s=0.06 (Study 1), r_s=0.17 (Study 2)**

PGA=physician global assessment; **=p not significant

Secondly, articles demonstrated relationships between stiffness and theoretically relevant constructs including disability, pain (Tables 2.6-2.7), and fatigue. A number of studies reported weak correlations between MS duration and disability while weak to strong correlations were reported between MS severity and disability (Table 2.6). However, some reported correlations included the whole study sample and not

exclusively RA patients (Wolfe, 1999; El Miedany *et al,* 2010; Jastrzabek *et al,* 2013) and disability assessment varied across studies. Overall, the evidence regarding the relationship between stiffness using current measures and disability is inconclusive. Although MS severity items appear to demonstrate stronger relationships with disability than MS duration, evidence is limited and comparisons are difficult given inconsistency and poor reporting of items.

Table 2.6: Correlations between stiffness items and disability

Authors	Stiffness item	Disability comparator	Results
Jastrzabek *et al,* 2013	Duration of MS	HAQ (Fries *et al,* 1980)	r_s=0.42, p=0.0068
El Miedany *et al,* 2010	Duration of MS	HAQ (Fries *et al,* 1980)	r_p=0.25, p<0.01
Khan *et al,* 2009	Duration of MS	HAQ (Fries *et al,* 1980)	r_s=0.43, p<0.001
Fransen *et al,* 2000	Duration of MS	HAQ (Fries *et al,* 1980)	r_s=0.37, p<0.0001
Vliet Vlieland *et al,* 1997	Duration of MS Severity of MS	HAQ (Fries *et al,* 1980) HAQ (Fries *et al,* 1980)	r_s=0.04** (Study 1), r_s=0.24, p<0.05 (Study 2) r_s=0.42, p<0.05 (Study 1), r_s=0.26, p<0.05 (Study 2)
Houssien, McKenna and Scott, 1997	Duration of MS	HAQ (Fries *et al,* 1980)	r_s=0.33, p<0.05
Westhoff *et al,* 2008	Severity of MS	Hannover functional questionnaire (Raspe *et al,* 1990, in Westhoff *et al,* 2008)	r_s=0.52, p<0.001 (baseline), r_s=0.58, p<0.001 (three year follow-up)
Wolfe, 1999	WOMAC stiffness indices	WOMAC functional indices (Bellamy *et al,* 1988)	r_s=0.76, p<0.0001
Yazici *et al,* 2004	MS duration	MHAQ (Pincus *et al,* 1983)	OR 6.89, 95% CI 3.82-12.4

**=p not significant

The relationship between stiffness and pain was reported in a number of studies. As seen with disability, pain has been reported to correlate weakly with MS duration but weak to strong correlations have been reported with MS severity (Table 2.7). As with disability, the evidence regarding the relationship between pain and stiffness using

current measures is varied depending on the items used. Again, stronger correlations are observed for MS severity than MS duration however, evidence is limited.

Table 2.7: Correlations between stiffness items and pain

Authors	Stiffness item	Results
Sarzi-Puttini, 2002	Duration of MS	r_S=0.43, p<0.001
Khan *et al,* 2009	Duration of MS	r_S=0.48, p<0.001
Fransen *et al,* 2000	Duration of MS	r_S=0.49, p<0.0001
Vliet Vlieland *et al,* 1997	Duration of MS	r_S=0.19** (Study 1), r_S=0.36, p<0.05 (Study 2)
	Severity of MS	r_S=0.48, p<0.001 (Study 1), r_S=0.47, p<0.001 (Study 2)
Houssien, McKenna and Scott, 1997	Duration of MS	r_S=0.41, p<0.05
Westhoff *et al,* 2008	Severity of MS	r_S=0.66, p<0.001 (baseline), r_S=0.76, p<0.001 (three year follow-up)
Wolfe, 1999	WOMAC stiffness index	r_S=0.73, p<0.0001
Yazici *et al,* 2004	Duration of MS	OR 1.44, 95% CI 1.32-1.58

**=p not significant

The relationship between stiffness and fatigue was infrequently reported. In studies with RA patients only, MS duration was significantly associated with fatigue in regression analyses (OR 1.28, 95% CI 1.19-1.39) (Yazici *et al,* 2004) and correlated weakly with fatigue (r_S=0.39, p<0.001) (Khan *et al,* 2009). In a study including RA, OA and FM patients, a moderate correlation was demonstrated between the WOMAC stiffness index and fatigue (r_S=0.52, p<0.0001) (Wolfe, 1999).

Finally, some studies provided evidence that stiffness items could discriminate between expected groups. Three studies reported stiffness items discriminating patients in relation to disease activity. One study compared three and six month changes in flare domains between patients in flare and not in flare. MS severity demonstrated one of the largest standardised mean differences (SMD) (SMD=1.17, 95% CI 0.78-1.55) of all assessed variables along with physician global assessment (SMD=1.31, 95% CI 1.07-1.55), pain VAS (SMD=1.30, 95% CI 1.06-1.54), and body pain (SMD=1.24, 95% CI 1.00-1.48). These were similar to DAS28 (SMD=1.26, 95% CI 1.00-1.52) which was collected for reference. However, the CI's for MS severity

were wide, suggested to be the result of less stiffness data than other variables (Lie *et al,* 2014). Another study reported that MS duration was significantly different among patients defined by different disease activity states (F(35226)=273.8, p<0.001). A ROC demonstrated that MS duration could moderately differentiate active from inactive disease (area under the curve=0.74, 95% CI 0.72-0.75) (Khan *et al,* 2009). In another study, MS severity demonstrated marginally better ability to discriminate between RA patients with active and inactive disease (sensitivity=85%, specificity=44%) than MS duration (sensitivity=78%, specificity=30%). However, disease activity definitions were based on physician judgement rather than standardised assessment. The study also reported that MS severity (sensitivity=72%, specificity=31%) and MS duration (sensitivity=74%, specificity=30%) were unable to discriminate patients with RA from those with non-inflammatory conditions (Hazes, Hayton, and Silman, 1993). A further study reported ROCs of the stiffness change scores for patients that had demonstrated improvement consistent with ACR criteria. Here MS severity appeared more sensitive to change than MS duration (area under the curve 0.77 and 0.70 respectively) (Vliet Vlieland *et al,* 1997). Although severity items appear to perform better than duration items, overall the evidence regarding discriminant validity is inconclusive and the use of different item formats limits comparison.

2.4.1.3.3.2 Other measurement properties

There was limited evidence in relation to other measurement properties including test-retest reliability, internal consistency, ability to detect change, floor and ceiling effects and interpretability. Two studies provided evidence regarding internal consistency (Leeb *et al,* 2003; Fransen *et al,* 2000). The SACRAH demonstrated good internal consistency overall (α=0.98) and for each domain (function α=0.98, stiffness α=0.79, pain α=0.90), although these results were generated from the whole sample which included patients with OA and RA. However, strong correlations (r_S=0.80-0.86) were reported between domain items and the total SACRAH specifically for RA patients (Leeb *et al,* 2003). Good internal consistency was also reported for the RADAI (α=0.87), which could have been increased slightly with the removal of the MS item (α=0.89). The stiffness item demonstrated moderate correlations with all combined RADAI items (r_S=0.51), although correlations between MS duration and each item individually were weak (disease activity last 6 months r_S=0.37, disease activity today r_S=0.46, pain r_S=0.47, tender joints r_S=0.48, all p<0.0001).

Four studies reported on ability to detect change (Vliet Vlieland *et al,* 1997; Buchbinder *et al,* 1995; Borstlap *et al,* 1995; Ward *et al,* 1994). One study demonstrated that stiffness scores significantly improved across time points (pre-operative, 3, 6, 12 months) in the RA group (Borstlap *et al,* 1995). However, this study provided no description of the stiffness question used other than that it was measured on a VAS. Another study reported that MS duration was not sensitive to changes in clinical status (Ward *et al,* 1994). However, sensitivity to change was not based on outcomes assessed before and after a treatment but over time, with estimates of clinical change based on minimum and maximum scores of other variables (physician global assessment, PtG, disability index, ESR). The time interval between the maximum and next minimum value represented the rate of change and this time interval was used to explore scores of other variables (such as stiffness). Thus the relationship between the variables used to generate the rate of change and the comparison variable could influence the results. Although a range of assessments (laboratory, clinician and patient) were used and MS duration performed consistently poorly across all measures. Another study investigated the ability of commonly used outcome measures to detect treatment effects. Here the relative efficiency (RE) of variables was compared to that of TJC. Although the RE of MS duration was only 0.23, it was not significantly different to TJC, as were pain (measured on a 5-point scale, RE=0.18) and ESR (RE=0.01). However, it was lower than other patient-reported variables including PtG (RE=1.17) and pain (measured on a 10cm VAS, RE=0.45) (Buchbinder *et al,* 1995). Finally, one article involving two studies investigated the ability to detect change using different stiffness items (Vliet Vlieland *et al,* 1997). The first study was observational and involved 63 RA patients who were hospitalised because of disease activity or functional decline. As expected, improvements in outcomes were reported between hospitalisation and discharge. When comparing different stiffness items, MS severity (effect size (ES)=0.74, standardised response mean (SRM)=0.64) demonstrated greater change than MS duration (ES=.41, SRM=.46). The second study was an RCT involving 80 RA patients starting or changing treatment. Differences between treatment and control group were greater for MS severity (two weeks ES=0.68, 12 weeks ES=0.49, 52 weeks ES=0.43) than MS duration (two weeks ES=0.30, 12 weeks ES=-0.07, 52 weeks ES=-0.16) at all time points, and were significantly different at the two later time points (12 weeks z=2.49, p=0.013 and 52 weeks z=2.60, p=0.009). Overall the article concluded that MS severity was a responsive outcome in comparison to other outcomes and MS duration (Vliet Vlieland *et al,* 1997).

2.4.1.4 Discussion

This review aimed to identify currently available stiffness PROMs and provide a summary of the evidence of their measurement properties. The 19 identified studies contained 37 individual stiffness assessment measures. All studies that defined the concept of stiffness assessment were based on either duration or severity and most assessed MS or EMS. However, despite the narrow focus of items, the variation in assessment of these concepts was considerable and item definition was poor, even in studies where stiffness assessment was the primary purpose. There was limited evidence of the measurement properties of stiffness items with evidence principally related to construct and criterion validity. However, given the variation and poor reporting of items, it was difficult to compare across studies, highlighting the difficulties posed when using stiffness PROMs in research.

Acknowledging the limited evidence and difficulties with item comparison, severity items appeared to perform better than duration items. Severity items displayed stronger construct and criterion validity and better discriminatory ability than MS duration. Of the six studies that contained assessment of both severity and duration items, four articles specifically recommended severity items over duration items based on their performance (Hazes, Hayton, and Silman, 1993; Vliet Vlieland *et al,* 1997; Westhoff *et al,* 2008; Lie *et al,* 2014). For example, Westhoff *et al.* (2008) stated that despite assessing both MS severity and MS duration, they only reported results for MS severity because it was more responsive. Conversely, in stiffness assessment in low-disease states, the two identified studies (Hazes, Hayton, and Silman, 1993; Khan *et al,* 2009) made conflicting recommendations regarding whether severity or duration was best (van Tuyl, Lems and Boers, 2014). Overall, there appears to be no clear evidence regarding the most appropriate measure to use to assess stiffness in RA.

From a content validity perspective, little evidence suggested that current stiffness measures have been developed using recommended methodology including qualitative exploration (USDHHS FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b). This may in part explain some difficulties identified with duration items. One study suggested that patients find completing duration items difficult and are often forced to report a cut-off time (Hazes *et al,* 1994). In another study, 19 participants reported that they had no stiffness when responding to an MS duration item yet reported a measurable amount of stiffness on an MS severity item (Vliet Vlieland *et al,* 1997). This may indicate difficulties for patients completing duration items or that

the items capture different information. Uncertainty regarding the content of duration items has also been reported by experts in RA and/or spondylarthropathy (Berthelot *et al,* 2002). Participants were asked "when you ask patients how long their morning stiffness lasts, do you indicate: until there is no more stiffness or until maximal improvement of stiffness is reached?" (p.149). There was inconsistency in responses, with six and 26 of the 32 responders indicating the respective options (Berthelot *et al,* 2002). In terms of severity, other studies have suggested that there are no difficulties for patients (Rhind, Unsworth and Haslock, 1987; Vliet Vlieland *et al,* 1997). These suggestions may relate to work indicating that MS severity has more impact on RA patients than MS duration (Mattila, Buttgereit and Tuominen, 2014), and that patients would pay more for a reduction in MS severity than MS duration (Tuominen, Tuominen and Möttönen, 2011; Tuominen, Tuominen and Möttönen, 2012). These results are particularly interesting when considering that duration items are most frequently implemented in research trials (Kalyoncu *et al,* 2009). It has been suggested that further research into different wording of stiffness items and better understanding of the value of assessing duration versus severity would be a useful addition to the literature (Lie *et al,* 2014). Further work therefore, to explore the patient perspective regarding stiffness assessment would be appropriate to enhance content validity and would also be consistent with PROM development recommendations (USDHHS FDA, 2009).

Having stated that the evidence regarding measurement properties provided by this review was limited, it is acknowledged that this review is not exhaustive. The search strategy used specifically identified articles containing measurement property information thus identifying a manageable number of articles containing relevant information. However, given that stiffness is a commonly used outcome measure (Kalyoncu *et al,* 2009), there will be many studies that were not identified by this review which may provide additional evidence for some or all of the stiffness assessment measures identified. This review was also limited in its focus on measures developed for an RA population. However, the search strategy was initially broad and included a wide range of rheumatic conditions so as to include any measures that may have been validated or tested in an RA population. Studies tested in an RA population were specifically included in Round 2 of the review and the broader literature including other conditions is addressed later (Section 2.4.2). Another limitation of the review is that there was no assessment of the quality of the identified studies. The strengths of the review include performing an update on a high quality, published systematic literature review (van Tuyl, Lems and Boers, 2014).

Overall, this review indicates that there is currently no clear evidence regarding the most appropriate measure to use to assess stiffness in RA. Current stiffness assessment relies on non-standardised and unvalidated EMS/MS duration or severity questions, which do not appear to have been developed according to current standards including collaboration with patients (USDHHS FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b), or methods recommended by OMERACT (Boers *et al,* 1998). Further work into the development of a stiffness measure with appropriate content validity would be a beneficial addition to the literature, as would further work to test defined stiffness measures to provide recommendations regarding the most appropriate stiffness tool to use in future research.

2.4.2 Identification of other stiffness measurement literature

Stiffness is not exclusively experienced by people with RA but also by those with other rheumatic diseases. Discussion at the OMERACT 12 conference (Orbai *et al,* 2015) highlighted that a number of PROMs used in other rheumatic conditions include stiffness items and that investigation into stiffness assessment in a broader rheumatology context would be relevant. As such, a scoping review was performed to identify common tools used in a rheumatology context that included stiffness items. The scoping review was performed based on expert suggestions during discussion at OMERACT and was furthered by evaluation of known articles, and other hand searching to identify relevant measures. In addition, some RA specific scales that were known to the researcher, and included stiffness items but had not been identified in the earlier review were also included here. The development and validation papers were identified and the measures used defined.

Table 2.8: Papers that describe scale development and/or validation where stiffness item/s are included within the scale

Population	Authors	Scale	Concept	Stem question	Response options/anchors	Timeframe
AS	Garrett *et al,* 1994	Bath ankylosing spondylitis disease activity index (BASDAI)*	Severity of MS	How would you describe the overall level of morning stiffness you have had from the time you wake up?	10cm VAS: 0 (None) to 10 (Very Severe)	Past week
			Duration of MS	How long does your morning stiffness last from the time you wake up?	10cm VAS: 0 (0 hours) to 10 (2 or more hours) with marked intervals at ½ hour, 1 hour, and 1½ hours	Past week
	Lukas *et al,* 2009	Ankylosing spondylitis disease activity score (ASDAS)*	Duration of MS	How long does your morning stiffness last from the time you wake up?	10cm VAS: 0 (0 hours) to 10 (2 or more hours) with marked intervals at ½ hour, 1 hour, and 1½ hours	Past week
FM	Burckhardt, Clark, Bennett, 1991	Fibromyalgia impact questionnaire (FIQ)	Severity	How bad has your stiffness been?	100mm anchored VAS: no stiffness to very stiff with marked increments	Past 7 days
	Bennett *et al,* 2009	Revised (FIQR)	Severity	Please rate your level of stiffness	100mm anchored VAS: no stiffness to severe stiffness with marked increments	Past 7 days
PMR	Leeb and Bird, 2004	PMR Activity Score*	Duration of MS	Exact wording unclear	Minutes	Unclear

Population	Authors	Scale	Concept	Stem question	Response options/anchors	Timeframe
OA	Leeb *et al,* 2003	Score for assessment and quantification of chronic rheumatic affections of the hands (SACRAH)	Daily MS severity	Exact wording unclear	100mm anchored VAS: no stiffness to unbearable stiffness	Unclear
			Starting stiffness after a time of rest	Exact wording unclear	100mm anchored VAS: no stiffness to unbearable stiffness	Unclear
	Sautner *et al,* 2004	Modified (M-SACRAH)	Daily MS severity	Exact wording unclear	100mm anchored VAS: no stiffness to unbearable stiffness	Unclear
			Starting stiffness after a time of rest	Exact wording unclear	100mm anchored VAS: no stiffness to unbearable stiffness	Unclear
	Rintelen *et al,* 2009	Short form (SF-SACRAH)	Severity of MS	How severe was your joint stiffness immediately after waking up first thing in the morning?	Anchored Likert scale: 0 (no stiffness) to 10 (unbearable stiffness)	Last 48 hours
	Bellamy *et al,* 1988	The Western Ontario and McMaster universities osteoarthritis index (WOMAC)	Severity of MS	How severe has your stiffness been after you first woke up in the morning?	Both validated in 5-point Likert scale (0=none, 1=mild, 2=moderate, 3=severe, 4=extreme), 100mm VAS (0 (none)-10 (extreme)), and 11-point NRS (0 (none)-10 (extreme)). Two items transformed to one score. Item wording begins with: Think	Last 48 hours
			Stiffness after rest	How severe has your stiffness been after sitting or lying down or while resting later in the day?		

Population	Authors	Scale	Concept	Stem question	Response options/anchors	Timeframe
OA cont.					about stiffness (not pain) you felt during the last 48 hours caused by the arthritis in your knee to be injected. Stiffness is the sensation of decreased ease in moving your joint.	
	Bellamy *et al,* 2002	Australian Canadian osteoarthritis hand index (AUSCAN)	Severity of MS	Targets stiffness after first wakening in the morning (exact wording unclear)	Validated in 5-point Likert scale (0=none, 1=slight, 2=moderate, 3= severe, 4=extreme), 100mm VAS (0 (none)-10 (extreme)), 11-point NRS (0 (none)-10 (extreme))	Last 48 hours
RA	Mason *et al,* 1992	Rapid assessment of disease activity in rheumatology (RADAR)*	Duration of MS	Were your joints stiff when you woke up today? No/Yes If yes, how long did this extra stiffness last?	7-point Likert scale (0=none, 1=<30 minutes, 2=30 minutes to an hour, 3=1-2 hours, 4=2-4 hours, 5=more than 4 hours but less than all day, 6=all day)	Today
	Stucki *et al,* 1995	Rheumatoid arthritis disease activity index (RADAI)*	Duration of MS	Were your joints stiff when you woke up today? No/Yes If yes, how long did this extra stiffness last?	7-point Likert scale (0=none, 1=<30 minutes, 2=30 minutes to an hour, 3=1-2 hours, 4=2-4 hours, 5=more than 4 hours but less than all day, 6=all day)	Today
	Leeb *et al,* 2008	RADAI-5*	Duration of MS	Did you experience joint (hand) stiffness on awakening yesterday morning? If yes, how long was this stiffness?	11-point Likert scale: 0 (no stiffness)-10 (stiffness the whole day)	Yesterday

Population	Authors	Scale	Concept	Stem question	Response options/anchors	Timeframe
RA cont.	Choy *et al,* 2008	Patient-Based Disease Activity Score Without ESR (PDAS2)*	Duration of MS	Were your joints stiff when you woke up today? No/Yes If yes, how long did this extra stiffness last?	7-point Likert scale (0=none, 1=<30 minutes, 2=30 minutes to an hour, 3=1-2 hours, 4=2-4 hours, 5=more than 4 hours but less than all day, 6=all day)	Today
	Anderson, 2001	Rheumatoid arthritis pain scale (RAPS)	Duration of MS	I have morning stiffness of one hour or more?	7-point Likert scale: 0 (always)-10 (never)	Last week
			Stiffness after rest	I feel stiffness in my joints after rest?	7-point Likert scale: 0 (always)-10 (never)	Last week

*disease activity composite assessment

The scoping review identified 15 scales containing 20 individual items (Table 2.8). As in the systematic literature review (Section 2.4.1), the identified stiffness items generally assessed the concepts of duration or severity. Most items assessed MS, although some assessed stiffness after rest (Bellamy *et al,* 1988; Anderson, 2001; Leeb *et al,* 2003; Sautner *et al,* 2004), and others assessed stiffness more broadly (Burckhardt, Clark, Bennett, 1991; Bennett *et al,* 2009). There was variation in stiffness item wording, response options, format, and timeframe. Although there was slightly better definition of items, this was still not consistent across all scales and identification of the precise question often entailed exploring a number of validation papers or scale documents.

There was replication of some items, as a number of the scales identified were updates of previous scales. For example, the original RADAR (Mason *et al,* 1992) was developed into the RADAI (Stucki *et al,* 1995), from which the PDAS2 (Choy *et al,* 2008) was developed, and all include the same MS duration item. However, no evidence was provided regarding the development of the content of the scale including involvement of patients. The RADAI-5 (Leeb *et al,* 2008) was developed to increase the ease of scoring of the original RADAI (Stucki *et al,* 1995) thus the anchors for all items were standardised on an 11-point Likert scale. However, the wording of the stiffness item was also changed to emphasise the hands and the timeframe was changed, yet no justification was provided (Leeb *et al,* 2008). The SACRAH (Leeb *et al,* 2003) was developed into the M-SACRAH (Sautner *et al,* 2004) and the SF-SACRAH (Rintelen *et al,* 2009). The SACRAH (Leeb *et al,* 2003) and M-SACRAH (Sautner *et al,* 2004) only define the question concept and the anchors used while the SF-SACRAH (Rintelen *et al,* 2009) defines the full item wording. However, the original SACRAH (Leeb *et al,* 2003) was developed in German and is not validated in English (Rintelen *et al,* 2009).

The WOMAC (Bellamy *et al,* 1988) and the AUSCAN (Bellamy *et al,* 2002) share similar items and development process. There is considerable evidence regarding the measurement properties of the WOMAC (McConnell, Kolopack and Davis, 2001) which was developed with substantial OA patient involvement (Bellamy and Buchanan, 1986). However, there is limited measurement property evidence for the stiffness subscale which demonstrates good internal consistency but has been reported to have inadequate test-retest evidence, and is omitted from some trials (McConnell, Kolopack and Davis, 2001). The AUSCAN was developed to assess hand pain, stiffness and disability in OA (Bellamy *et al,* 2002) and was based on the

rigorous item generation process used in the WOMAC (Bellamy *et al,* 1988). This involved item generation from a systematic literature review and clinician involvement in closed-question development for use in patient interviews that focused on the importance of the developed items (Bellamy *et al,* 2002). Therefore, although patients were involved in item review, they were not directly involved in item development (Poole, 2011).

The BASDAI contains two stiffness items, on MS severity and duration (Garrett *et al,* 1994). The BASDAI was developed with input from clinical and patient AS experts. However, no detail about the specific patient involvement was provided. The ASDAS was developed as a new method of disease activity assessment in AS (Lukas *et al,* 2009) and includes the MS duration question from the BASDAI (Garrett *et al,* 1994). The development of the ASDAS was based on the process used in the development of the DAS (van der Heije, 1993). A Delphi exercise involving clinical and patient AS experts was performed and after three Delphi rounds (where items were retained if endorsed by >80% responders), the MS duration item was one of 12 retained items. The advantage of this process was the involvement of patients in the Delphi. However, it was unclear whether both stiffness items from the original BASDAI were included in the Delphi exercise, and the voting process during each round was not reported, so there is no evidence as to why the duration item was selected and the severity item was not.

Finally the FIQ (Burckhardt, Clark, Bennett, 1991) and the updated FIQR (Bennett *et al,* 2009) were developed for FM assessment. The updated FIQR items were developed to overcome scoring difficulties with the original FIQ that restricted its use. Items were modified based on the original FIQ and relevant literature and were discussed with a focus group of 10 FM patients (Bennett *et al,* 2009).

The two remaining items were not based on development of other scales. The PMR activity score was developed as a composite assessment of disease activity (Leeb and Bird, 2004). The RAPS (Anderson, 2001) was developed specifically to assess pain in people with RA. Items were developed based on the content from pain theories within the literature and interviews with RA patients. However, as in many of the above scales, the involvement of patients in the development of items was not described.

This scoping review has identified that within the broader rheumatology literature, there are a number of validated questionnaires that include stiffness items. Although the validated nature of these tools provides evidence for their use, they were developed in populations other than RA. It is not known whether stiffness assessment could be general across conditions or whether disease specific stiffness assessment is required. The advantage of consistent assessment across diseases would be the ability for comparison. However, not enough is known about the consistency of the patient experience of stiffness across conditions to understand this fully. As such, an OMERACT stiffness special interest group has been endorsed to enable investigation across conditions (Orbai *et al,* 2015). In addition, although validated, the identified tools still demonstrate considerable variation in stiffness item wording, response options, format, and timeframe, and poor definition in some articles. Furthermore the measurement property evidence of stiffness items specifically is limited (e.g. WOMAC stiffness subscale), and as highlighted above, some of the concepts these items capture have been challenged in the RA specific literature (Section 2.4.1.4).

This review also highlights that there is very little documented evidence of the development process of these items, particularly in relation to patient involvement and the enhancement of content validity, which is an essential part of current PROM development guidelines (USDHHS FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b). Patient involvement in some scale development was poorly described (e.g. BASDAI, Garrett *et al,* 1994), limited to clinicians (e.g. SACRAH, Leeb *et al,* 2003), or to patients responding to closed-questions created by clinicians (AUSCAN, Bellamy *et al,* 2002). Furthermore, although the WOMAC (Bellamy *et al,* 1988) described a rigorous item development process involving patients, this was performed in an OA population and given the uncertainty regarding the general or specific nature of stiffness assessment, further work in an RA population is necessary to develop understanding.

It is acknowledged that given the review approach, there are likely to be other scales that include stiffness items within the literature that have not been identified, although the most common scales will have been recognised. Overall, consistent with the systematic literature review (Section 2.4.1), this scoping review suggests that despite the development and validation of a number of measures that include stiffness items, there is still considerable variation in the content of stiffness PROMs, poor definition of items, and little reported or implemented patient involvement. This highlights the need for further work into the development of a stiffness measure with appropriate

content validity, and identification of the most appropriate stiffness tool to use in stiffness assessment.

2.5 Importance of appropriate stiffness assessment

As stated earlier, there is currently no standardised method of assessing patient-reported stiffness in RA and no clear evidence regarding the most appropriate measure to use. Standardised assessment is vital in both research and clinical contexts to enable comparison across studies and consistent measurement of disease progress incorporating the patient perspective. The use of PROMs is vital when assessing concepts that are best understood by patients (USDHHS FDA, 2009). However, as identified earlier (Section 2.4.1), current stiffness assessment does not appear to have been developed according to current standards including collaboration with patients (USDHHS FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b), or methods recommended by OMERACT (Boers *et al,* 1998). Furthermore, there may be difficulties with some concepts currently used in stiffness assessment from the patient perspective. Therefore further work to explore the patient perspective of stiffness would fit with recommendations in the literature (e.g. Lie *et al,* 2014) and be relevant in the development of a RA stiffness PROM with appropriate content validity. This requires understanding how patients experience, conceptualise and evaluate stiffness so that questions to capture its essence can be developed, using language that patients understand (USDHHS FDA, 2009). The development of any new RA stiffness PROM should consider measurement properties, the OMERACT Filter (Boers *et al,* 1998), PROM development guidelines (e.g. Patrick *et al,* 2011a), and appropriate test theory (Section 2.3). Further work into testing any new RA stiffness PROM against current defined stiffness measures would also be important to provide recommendations regarding the most appropriate stiffness tool to use in future work, thus providing standardised stiffness assessment.

Chapter 3: Purpose and structure of research

The previous two chapters have discussed the relevance and assessment of stiffness within RA and other conditions, and considered PROM development. This chapter will describe the purpose and aims of this research, how they will be achieved and structured, the researcher perspective and aspects of research design.

3.1 Purpose of research

The purpose of this research is to better understand stiffness in people with RA, and then use this to develop and test a new RA stiffness PROM. The purpose of the development of a new RA stiffness PROM is to capture and assess the patient experience of stiffness in a standardised way. This is important because stiffness is a relevant patient symptom used in clinical and research settings (Chapter 1), but current stiffness PROMs are poorly defined and have limited measurement property evidence (Chapter 2). As PROM provide an assessment of a patient's health condition that comes directly from a patient, they are particularly useful when measuring concepts that are best understood by patients, such as disease symptoms (USDHHS FDA, 2009). Therefore, the development of a new stiffness PROM will provide standardised assessment that captures this patient relevant symptom. It is proposed that this will be developed for use in clinical and research environments.

The development of a new RA stiffness PROM will create the potential for stiffness to be included in the ACR disease activity core set, from which it is currently omitted because it was not sensitive to change (Felson *et al,* 1993). It would also address the OMERACT research agenda for development of a stiffness PROM in relation to flare (Bingham *et al,* 2011, Bykerk *et al,* 2014a). Most importantly, it aims to provide a standardised method of assessing a symptom that is important and relevant to patients, in research and clinical situations. This work may also lead onto further research into the assessment of stiffness across conditions.

3.1.1 Research aims

The overall purpose of this research is to explore the experience of stiffness in people with RA and use this to develop and test an RA stiffness PROM. The aims and objectives for each study can be found in subsequent chapters. The specific thesis objectives are detailed below:

- To understand the experience of stiffness in people with RA

- To explore which aspects from the RA patient experience of stiffness might be relevant in the patient-reported assessment of stiffness
- To develop a set of items that capture those patient-relevant aspects using appropriate wording and formatting
- To explore the acceptability of the draft items with people with RA
- To explore the performance of these items to develop the smallest and internally consistent set of items to form an RA stiffness PROM
- To test how these items perform compared to current stiffness assessment
- To make recommendations about the most appropriate way to assess stiffness in clinical and research environments

3.2 Researcher perspective

3.2.1 Prior knowledge

The researcher came to this project with a background in sport, health and physical activity. She had a particular interest in exercise and physical activity in long-term conditions and some experience of conducting small research projects. In the early stages of this PhD, focus was put into developing a better understanding of RA. This was achieved by reading within the rheumatology literature, attending rheumatologist and specialist nurse clinics within the department, and listening to the personal experiences of the patient research partners in the supervisory team. Engagement in research training courses provided by the University and the local hospital trust helped develop research skills and ensured compliance with good practice guidelines.

3.2.2 Epistemological position

The researcher's position for this research was based on pragmatism. Pragmatism encourages a focus on the research question and outcome (Creswell and Clark, 2011; Tashakkori and Teddlie, 1998). This emphasis enables freedom to choose methods or procedures most appropriate and compatible with the purposes of the research (Creswell, 2003, p.12) rather than being driven by an epistemological or ontological standpoint (Johnson and Onwegbuzie, 2004). As this research was driven by specific research questions and outcomes, freedom in relation to the choice of methods was vital in the planning and execution of each study. Specifically, a mixed methods approach was essential for this research and a philosophical underpinning of pragmatism fits well with such an approach (Denscombe, 2008).

3.3 Methodological approach

3.3.1 Mixed methods

As a result of considering the research objectives and outcome, it was clear that the use of both qualitative and quantitative methods would be necessary to answer the research objectives effectively. Mixed methods research is referred to as the third paradigm in addition to qualitative and quantitative research (Johnson, Onwuegbuzie and Turner, 2007). Although there are different opinions among researchers as to what constitutes mixed methods research (Sandelowski, 2000; Bryman, 2007; Tashakkori and Creswell, 2007), an overview of the literature by Denscome (2008) suggested that the characteristics of a mixed methods approach include: the use of both qualitative and quantitative methods within the same research project; clear description of the sequencing and priority given to each of these aspects, and pragmatism as the philosophical underpinning (Denscombe, 2008). Mixed methods research is used for a number of purposes including instrument development (Bryman, 2006), making it an appropriate choice for this research. This is also consistent with best practice guidelines for PROM development which recommend the use of both qualitative and quantitative methods (Fitzpatrick *et al,* 1998; USDHHS FDA, 2009).

3.3.1.1 Mixed methods sequencing

Within the literature there has been a drive for the development of classification systems of mixed methods research designs (e.g. Leech and Onwegbuzie, 2009; Creswell and Clark, 2011). Creswell and Clark (2011) propose a sequential exploratory strategy with an instrument development variant. This fits with the aims of this research as it involves initial exploratory qualitative phases, which inform the development of the draft items for quantitative testing.

3.3.2 Research design

This research used three phases to meet its objectives (Figure 3.1). The exploratory phase aimed to better understand the RA patient experience of stiffness and used qualitative interviews (Study 1, Chapter 4). The development phase then focused on the development of the content (items) for an RA stiffness PROM. Firstly, focus groups were used to validate the findings from Study 1 in a new sample of patients, and to explore stiffness specifically from a measurement perspective (Study 2, Chapter 5). Following this, draft items were developed in iterative rounds of discussion with the supervisory team and patient research partners (Chapter 6). Items

were then tested and refined using cognitive interviews (Study 3, Chapter 7). The final testing phase aimed to develop and subsequently test the item structure using a quantitative survey (Study 4, Chapters 8 and 9).

3.3.3 Thesis structure

This thesis is structured around nine chapters. It aims to capture the process of development and progress from one study to the next. The first three chapters contain background information and reviews of the literature. Chapters 4, 5, 7, 8 and 9 describe each of the four studies within the research (Figure 3.1) and each contain background, methods, results, discussion, and conclusion sections. Chapter 6 ensures a transparent process of item development. The final chapter is a discussion, consolidating the research findings and conclusions, and identifying areas for future research.

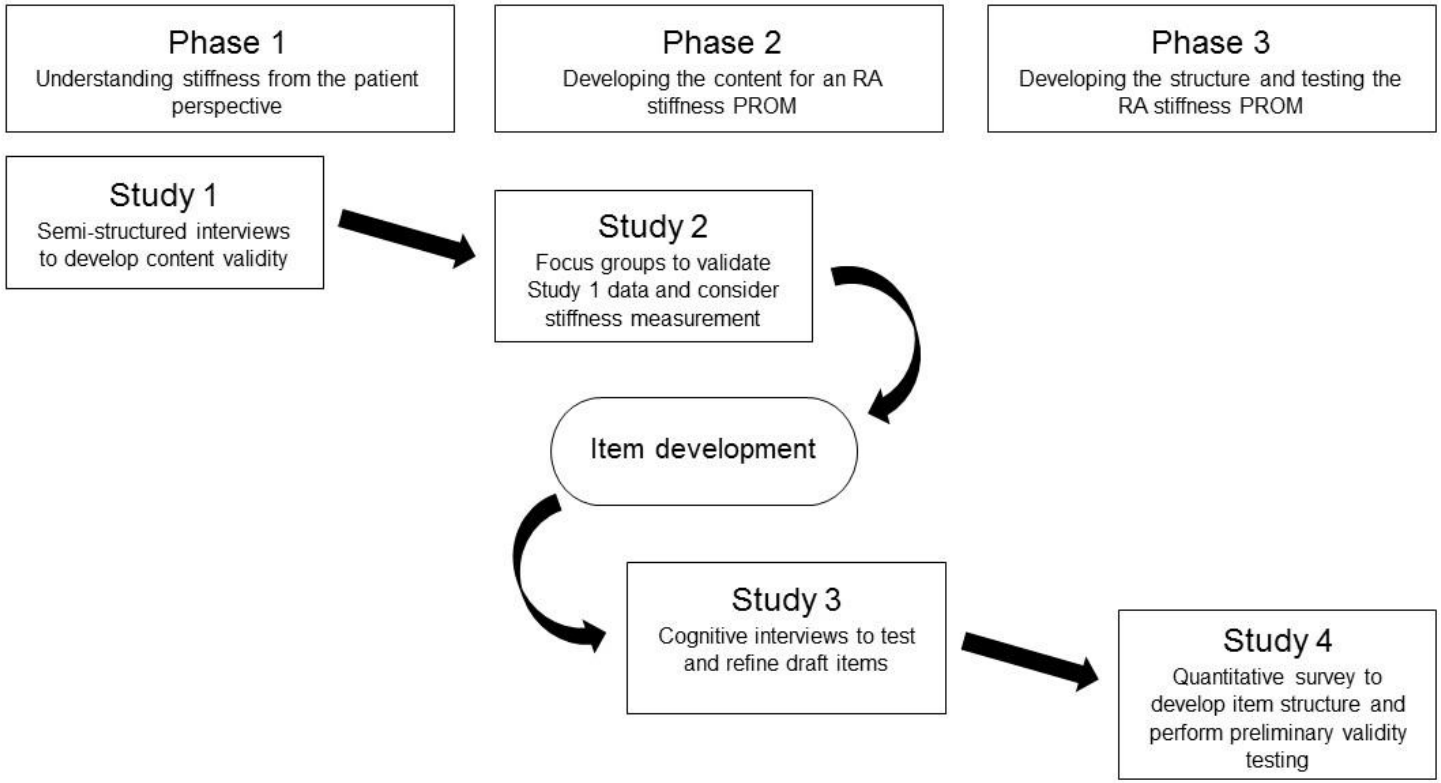


Figure 3.1: PhD outline

3.4 Research process and influences

In addition to the researcher's perspective and the methodological approach, there have been a number of other influences on the design and conduct of the research.

3.4.1 Patient research partner involvement

The public involvement national advisory group INVOLVE defines public involvement as "research being carried out 'with' or 'by' members of the public rather than 'to', 'about' or 'for' them" (INVOLVE, 2012). Within the literature, there is evidence for the involvement of the public and patients in research (Brett, 2010). In a recent literature review, the benefits of public involvement in a number of key areas of the research process were reported such as participant recruitment and project design (Stanley, 2009). A specific example of the benefits of public involvement is OMERACT which has been involving patients in research for 10 years and allocates 10% of its conference places to patients. A recent review aimed to explore the effect of patient involvement on OMERACT by reviewing conference documents and interviewing conference participants (de Witt *et al,* 2014). The review identified a number of facilitators and barriers to patient involvement which led to recommendations to enhance development of future patient involvement. A particular benefit of patient involvement was the identification of important areas for the research agenda, such as fatigue (de Witt *et al,* 2014). It also highlighted challenges to patient involvement in research, for which there are guidelines describing appropriate considerations when involving patients in research (e.g. Hewlett *et al,* 2006; de Witt *et al,* 2011; INVOLVE, 2012).

Public involvement has been suggested to be particularly important in qualitative research involving sharing views or experiences (Stanley, 2009), and the involvement of patients in PRO development has been reported to enhance relevance, acceptability and quality (Stanley, 2009; Staniszewska *et al,* 2012). Therefore, for this thesis it was vital not only to involve patients as participants but also to involve patients within the research process. Therefore from the outset of this project and throughout, two patient partners (GB and AE) were part of the supervisory team.

3.4.2 Supervisory team

Supervisory teams are made up of experienced individuals with knowledge and expertise in the topic of the research project. Table 3.1 provides a brief description of each team member, and the abbreviation that is used to identify them. The team has

a wide range of experience from academic, clinical, and experiential perspectives which have been vital throughout the planning and execution of this project.

Table 3.1 Supervisory team characteristics

Team	Gender	Position	Years of rheumatology experience
ED	F	Rheumatology psychology researcher (Senior Research Fellow)	7 years
JK	M	Academic rheumatologist (Emeritus Professor)	>30 years
JP	M	Epidemiologist (Associate Professor)	>10 years
GB	F	Patient research partner	RA diagnosed ≥10 years
AE	F	Patient research partner	RA diagnosed ≤10 years
SH	F	Academic rheumatology nurse (Professor)	>20 years

3.4.3 Reflexivity

Reflexivity has been defined as “thoughtful, self-conscious awareness” (Finlay, 2002, p.532). It is an acknowledgement of personal (e.g. gender) and intellectual (e.g. professional background) biases that may have influenced the research (Mays and Pope, 2000), and is used to enhance rigor in qualitative research (Koch and Harrington, 1998). However, there are criticisms of the practicalities of reflexivity, such as the suggestion that researchers need to have “superhuman self-consciousness” and have no problem accessing their feelings or motivations (Seale, 1999, p.168). Finlay (1998; 2002) argues that reflexivity is more of a resource than a problem and should be used as a research tool. In this context, the aim is not to achieve impartiality, but to use reflexivity tools to better understand the influence of bias (Frank, 1997). It is suggested that researchers state any biases (e.g. intellectual or personal) at the start of their research to enhance credibility of their findings (Goodwin, 2006). This includes disclosing professional background to readers and participants (Richards and Emslie, 2000). Initial statements of the biases such as background and perspective of the researcher and supervisory team have been presented (Sections 3.2.1 and 3.4.2). In relevant studies, professional background disclosure to participants has been described and a personal reflection of the research is given in Chapter 10.

The next chapter will describe the first study within this thesis aiming to understand the RA patient experience of stiffness.

Chapter 4: Understanding stiffness from the patient perspective (Study 1)

The previous chapters discussed the background literature surrounding RA, stiffness, and outcome measures, and outlined the aims and structure of this thesis. This chapter describes Study 1 which aimed to better understand stiffness from the patient perspective.

4.1 Background

Stiffness is commonly reported by RA patients (Scott, 1960; Vliet Vlieland *et al,* 1997; Khan *et al,* 2009) and it is known to affect patients' daily life, work life, and quality of life (Westhoff *et al,* 2008; da Silva *et al,* 2011; Phillips and Dow, 2012). It has also been demonstrated to be relevant to patients in relation to fluctuating aspects of the disease such as flare (Bartlett *et al,* 2012) and low disease activity (van Tuyl *et al,* 2015). Stiffness initially featured in early RA disease activity indices (e.g. the Lansbury index, Lansbury, 1958) and was included in the original RA classification (Arnett *et al,* 1988) and remission criteria (Pinals *et al,* 1981). Stiffness is commonly used as an outcome in research (Kalyoncu *et al,* 2009) and has been suggested to influence clinical decision making (Kirwan *et al,* 1984; Soubrier *et al,* 2006).

Despite the broad relevance and uses of stiffness, it remains poorly understood and inconsistently assessed. Current stiffness assessment is based on EMS/MS duration or severity items which are often poorly defined. There is no clear evidence regarding the most appropriate measure to use to assess stiffness in RA, and no standardised approach (van Tuyl, Lems and Boers, 2014). There are reported difficulties with current assessment methods (e.g. Vliet Vlieland *et al,* 1997; Westhoff *et al,* 2008), and items do not appear to have been developed according to current standards (USDHHS FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b). The one previous study that focused on understanding the patient experience of stiffness (Lineker *et al,* 1999) was performed over a decade ago, since when there have been substantial changes in RA treatment (Smolen *et al,* 2010) and therefore possibly changes in stiffness experiences. In order to work towards better assessment of stiffness, understanding the patient experience is essential. This will enable development of a stiffness measure with appropriate content validity.

4.2 Aims and objectives

The overall aim of this study was to develop a better understanding of the experience of stiffness in people with RA. The specific study objectives were:

- To investigate the experience of stiffness and how it is described by people with RA
- To understand how people with RA evaluate stiffness and how they describe changes in it
- To understand how people with RA describe stiffness in relation to other symptoms such as pain

4.3 Methods

4.3.1 Semi-structured interviews

Qualitative methods involve the use of words as data (Braun and Clarke, 2012) and are a way of exploring and understanding the views and experiences of participants (Mays and Pope, 1995). Qualitative methods can encompass a variety of data collection approaches including interviews and focus groups (Braun and Clarke, 2012). Semi-structured interviews are a method of investigating a particular topic using a flexible structure of open-ended questions. From these open-ended questions the interviewer can follow-up responses in more detail with further questions (Britten, 2006). Semi-structured interviews are useful because they provide structure but also flexibility; as such, topics can be discussed in the order most appropriate to the participant, enable further detail to be given around topics of discussion, and provide an opportunity for unexpected ideas to be generated and discussed (Arthur and Nazroo, 2003; Britten, 2006). The study aims and objectives could have been achieved with other qualitative methods such as focus groups. Focus groups naturally facilitate interaction between group participants and are useful to explore and clarify the views of the group on particular topics (Kitzinger, 1994). However, it has been suggested that focus groups can be negatively influenced by dominant participants and provide a less detailed understanding of a topic than other methods such as semi-structured interviews (Krueger and Casey, 2009). Therefore, to enable a detailed exploration of individual patients' experiences of stiffness it was considered that semi-structured interviews would be the most appropriate method for this study.

4.3.2 Participant identification and sampling

Ethics approval was obtained from the University of the West of England research ethics committee (REC) (HLS/13/01/26) and from the Leeds East REC following

proportionate review (13YH0050). The recruitment criteria defined that participants had a confirmed diagnosis of RA (Arnett *et al,* 1988; Aletaha *et al,* 2010), were aged 18 years or over, could speak English to a sufficient degree to participate in the study unaided, and had self-defined experience of RA related stiffness. Qualitative research commonly utilises purposive sampling where individuals with the relevant experience or insight are recruited to provide information on the topic (Patton, 2002). Within the purposive sampling approach, a mix of age, gender and disease duration were targeted to ensure that a range of participants were recruited (Sandelowski, 1995; 2000). Data saturation is commonly used as a guideline regarding sample size however, sample size recommendations in the literature vary and are often not supported by rationale (Guest, Bunce and Johnson, 2006). Data saturation, defined as the point at which no new information is generated (Guest, Bunce and Johnson, 2006) has been criticised for having multiple interpretations (O'Reilly and Parker, 2013). In an attempt to identify a target sample for this study, it has been suggested that researchers consider other studies using similar designs where saturation was achieved when deciding on adequate sample size (Onwuegbuzie and Leech, 2007). As such, following a review of the literature it was decided that 15-20 participants should be initially recruited. Recruitment could then be continued if required.

Recruitment took place at hospital out-patient rheumatology clinics at the Bristol Royal Infirmary (BRI) in University Hospitals Bristol Trust and at Cossham Hospital in North Bristol Trust (NBT). At both sites, interested potential participants were given a patient information sheet (PIS, Appendix B) to take away and consider. A reply slip, prepaid envelope and contact information were provided so that interested participants could contact the researcher by post, telephone or email to discuss the study further and arrange a convenient date and time for an interview.

4.3.2.1 Site specific differences in participant identification and approach

There were slight differences in the identification and approach of potential participants at the two recruitment sites. At the BRI, eligible participants were approached directly by the researcher while in the waiting room. At NBT, eligible participants were identified and initially approached by a member of their direct clinical team. If the eligible potential participant agreed to hear more about the study, they were introduced to the researcher to discuss the study further in a private clinic room.

4.3.3 Interview guide development

An interview guide (Table 4.1 and Appendix C) was developed to explore the aims of the study and was based on a literature review and discussion with the research team, particularly the patient research partners (GB and AE). The interview guide was flexibly observed during each interview, as such, the first question "Can you tell me about your experience of stiffness in relation to RA?" was asked and subsequent questions and prompts followed depending on each individual participant's response. Interviews followed an iterative process, which allowed ideas and concepts identified in early analysis to be explored in subsequent interviews (Legard, Keegan and Ward, 2003). In addition, participants were asked to describe stiffness using the metaphor of an animal. Metaphors are commonly used in discourse and have been utilised in healthcare research to describe personal experiences of diseases (Youngson *et al,* 2015) and symptoms (e.g. Wylde *et al,* 2014). They are also specifically used within a rheumatology context, for example in Sjögren's syndrome where 'gritty eyes' is often used in patient literature (Arthritis Research UK, 2014), and the use of a metaphor of an animal has been used in previous work with RA patients (Flurey, 2012). Metaphors are particularly useful with concepts that are difficult to describe literally (Ortony, 1975) thus this question attempted to capture the potentially abstract concept of stiffness.

Table 4.1: Interview guide

A. Can you tell me about your experience of stiffness in relation to RA?
B. How does this vary in a 24 hour period?
C. Has stiffness changed over the course of your disease?
D. How does stiffness differ from other RA symptoms?
E. What are the consequences of stiffness?
F. How do you deal with stiffness?
G. How to you assess stiffness?
H. Is there anything that you feel is important to stiffness that we have not talked about?
I. If your stiffness was an animal what would it be and why?

A pilot interview was performed with one patient research partner (GB) and was observed by a member of the research team (ED) with qualitative research expertise. This provided an opportunity to test and refine the interview guide and allowed the researcher to gain experience and confidence prior to commencing interviews with participants.

4.3.4 Interview procedure

Eighteen participants agreed to participate. However, one individual did not wish to be audio recorded and following consideration it was decided that this would produce inconsistent data. This was explained to the individual who was thanked for her interest. Another individual cancelled our first interview appointment and despite original enthusiasm to participate was unable to rearrange the appointment around her busy full-time work schedule. Other reasons given for declining participation included recent participation in other research studies and time commitments such as work, children or hospital appointments.

Interviews took place in non-clinical rooms in the Academic Rheumatology Unit at the BRI and at Cossham Hospital. All interviews were performed by one researcher who was unknown to participants prior to the study and introduced herself as a doctoral student with a non-clinical background. Most participants (n=13) chose to attend an interview at the location in which they normally attended clinic. Three participants chose to be interviewed at the other site for convenience and were provided with maps and directions if required. Each participant was greeted by the researcher and was provided with refreshments. Prior to commencing each interview, participants gave informed consent, completed a questionnaire pack (Section 4.3.4.1) and were asked if they had any questions. All interviews lasted between 30-80 minutes and were conducted with only the researcher and participant present, apart from one interview where the participant was accompanied by her young son. Interviews were audio-recorded and transcribed verbatim. Two interview recordings were transcribed by the researcher. All other recordings were transcribed by a transcription service but were checked for accuracy and anonymised by the researcher.

4.3.4.1 Questionnaire pack

A short questionnaire pack was developed to describe the recruited sample and their perceived level of disease severity and disability. The pack (Appendix D) contained demographic, clinical disease measures and medication questions. A brief description of each of the validated items is given below.

4.3.4.1.1 Health assessment questionnaire (HAQ)

The HAQ is a 20 item patient report of functional disability that focuses on eight categories (dressing and grooming, rising, eating, walking, hygiene, reach, grip, activities) (Fries *et al,* 1980). Patients rate questions within each section with a score

between 0 (without any difficulty) and 3 (unable to do). The highest score from each section is then added together to produce an overall score. This overall score is then averaged to give a total score between 0 and 3 where higher total scores indicate worse perceived function and disability.

4.3.4.1.2 PtG

The PtG assessment of general health is a 10cm VAS which asks patients to indicate how well they are doing with their arthritis (0=very well, 10=very badly). The PtG is one of two patient reports in the validated DAS28 (van der Heijde *et al,* 1993).

4.3.4.1.3 Pain

This assessment of pain is a 10cm VAS which asks patients how much pain they have experienced as a result of their arthritis within the last week (0=no pain, 10=severe pain) (Farrah *et al,* 2001; Hawker *et al,* 2011).

4.4 Analysis

Thematic analysis is a method of identifying, analysing and reporting patterns in collected data (Braun and Clarke, 2006). A justification for thematic analysis being appropriate for this study is provided below.

4.4.1 Analysis approach

Thematic analysis (Braun and Clarke, 2006; 2012) was considered most appropriate for use in this research, although three approaches to analysis were considered; interpretive phenomenological analysis (IPA); grounded theory (GT); and thematic analysis. IPA allows exploration into the lived experiences of particular phenomena of individuals or small groups of people (Smith, Flowers and Larkin, 2009) which made it a sensible consideration for use in this study. IPA has been praised for its accessibility for those unfamiliar with the approach given its detailed guidance (Smith, Flowers and Larkin, 2009), and for its suitability in research projects with time and resource limitations (Braun and Clarke, 2012). However, there are criticisms of IPA including that it is labelled by some as a 'descriptive' approach (Larkin, Watts and Clifton, 2006, p.102), and its precise guidelines allow little flexibility (Braun and Clarke, 2012). It has also been suggested that because IPA focuses on individual participants as well as patterns across participants that this dual focus may not enable the same detail as provided by approaches such as thematic analysis or GT (Braun and Clarke, 2012). GT aims to generate plausible and useful theory that is grounded

in the data (Glaser and Strauss, 1967; McLeod, 2001). This 'bottom up' or 'data driven' approach (McLeod, 2001) made it a sensible consideration for this study, especially considering its focus on the development of theory (Holloway and Todres, 2003). However, GT has been criticised for being time consuming, and suggested as most effective in large research projects (Braun and Clarke, 2012) which from a practical perspective, was not compatible with a small scale PhD project. GT also has multiple versions (e.g. Charmaz, 1990), which can create uncertainty about the most appropriate version to utilise (Birks and Mills, 2011). One of these versions, termed GT 'lite', encourages a GT approach without the strong emphasis on the theoretical position. However, it has been argued that most GT approaches actually are GT 'lite' and that this approach is very similar to thematic analysis (Braun and Clarke, 2006). Within the literature, thematic analysis can include a number of approaches, such as Boyatzis' (1998) 'process' that can be used with qualitative information, and Attride-Stirling's (2001) visual thematic networks, which vary in the approach taken and outcome attained. However, Braun and Clarke's thematic analysis (2006; 2012) provided a standardised name and approach to performing thematic analysis. Thus subsequent discussion of thematic analysis refers to thematic analysis as described by Braun and Clarke (2006; 2012). Thematic analysis presents clear guidelines for performing the method and provides an opportunity to learn basic data handling strategies (Braun and Clarke 2006; 2012) which was considered beneficial for a developing qualitative researcher. Additionally, thematic analysis provided a flexible approach to analysis as it is a method only approach and does not enforce other aspects such as data collection methods or theoretical positions (Braun and Clarke, 2012) as seen in other approaches such as GT. It is also flexible in that it is applicable in a 'bottom up' (inductive) and 'top down' (deductive) manner (Braun and Clarke, 2012). Inductive thematic analysis involves coding in a data driven manner while deductive thematic analysis is coded in an approach driven by pre-existing theory (Braun and Clarke, 2006). The 'data driven', inductive thematic analysis was most appropriate for Study 1 to develop understanding of a patient symptom. Overall, the clear guidance and flexible nature of thematic analysis were key factors in its consideration as appropriate for use in this research.

4.4.1.1 Thematic analysis process

Thematic analysis involved transcripts being read, re-read, and systematically coded, codes were then explored for patterns, which led to theme development. The six stages of this process have been described below and include description of specific actions and reflexive notes from each stage for transparency.

4.4.1.1.1 Stage 1: Familiarisation

The first stage of the thematic analysis process involves becoming immersed in the data, allowing the researcher to detect features that are relevant to the research question. To familiarise herself with the data the researcher transcribed two audio recordings, and checked the accuracy of all other transcripts by listening to the audio-recording while also reading each transcript. Early thoughts and ideas were noted by hand on the transcripts to begin understanding the data.

4.4.1.1.2 Stage 2: Coding

Coding is a process of identifying and labelling features of data that are relevant to the research question. Although this process was conducted systematically, as the researcher used software with which she was unfamiliar, the approach was initially experimental and led to checking data across different software packages. This was likely to have been time inefficient however, all data were treated equally and this approach did mean that data were rigorously checked. First, each transcript was coded by hand by highlighting relevant text and labelling it. The researcher then used NVivo 10 (QSR International Pty Ltd, 2012) to broadly 'bucket code' features relating to similar topics within each transcript. Following this, NVivo 10 (QSR International Pty Ltd, 2012) and Microsoft Office Word 2013 were used to code data again by highlighting and labelling relevant text. This was double checked against the 'bucket codes'. As the coding structure developed it was reviewed and previously coded transcripts were also revisited and re-coded. Two transcripts were also independently coded by two members of the supervisory team (SH, ED) with experience of qualitative analysis. Following a brief introduction, patient research partners (GB, AE) also read two transcripts and highlighted relevant points from their perspective. Following the completion of five interviews, a team meeting was held to discuss early analysis and possible directions for future interviews.

4.4.1.1.3 Stage 3: Searching for themes

The process of identifying patterns or themes involves reviewing the codes and trying to identify connections. Originally, the researcher worked around the ideas generated for the original 'bucket codes' as themes. She revised and reworked topics that related to group of codes. However, Braun and Clarke (2012) also suggest that a theme has a central organising concept that should say something meaningful about the data. The researcher found it difficult to derive a clear central organising concept around these themes that also encompassed the patient voice. So at this point the

researcher went back to working by hand using diagrams and the 'long table approach' (Krueger and Casey, 2009) to group codes to try and explain these data more clearly. This was refined into a thematic diagram to visualise relationships between themes.

4.4.1.1.4 Stage 4: Reviewing themes

The next step of analysis has been described as a 'quality control' measure involving checking that the candidate themes in the analysis fit with the overall dataset. Here the researcher reviewed the thematic diagram and coding structure to ensure that nothing major had been omitted. A suggested thematic diagram was taken to a team meeting where the proposed themes were presented and the diagram was refined with the research team and the patient research partners. Once the thematic diagram was finalised, each of the transcripts was re-read to ensure that the key ideas were correctly captured.

4.4.1.1.5 Stage 5: Defining and naming themes

Writing theme definitions is an essential part of being able to define each theme. Theme names were derived iteratively during team meetings and in particular during discussion with ED. This really aided capturing the data from the patient perspective and highlighted the patient voice in the analysis.

4.4.1.1.6 Stage 6: Writing – finalising the analysis

Braun and Clarke (2012) argue that analysis is not complete until it is written, as writing helps clearly define the themes. A draft chapter was written which was refined and published as a journal article (Halls *et al,* 2015).

4.5 Results

4.5.1 Participants

Sixteen of the 38 individuals approached agreed to participate (42% recruitment rate): 11 were female (69%), age range between 33 and 78 years and disease duration between 1 and 27 years (Table 4.2).

Table 4.2: Participant demographic information

Pt ID	Gender	Age (Yrs) †	Dis dur (Yrs) ‡	NHS site	HAQ §	PtG ¥	Pt pain ¤	Current medication	Work status	Education
101	M	62	22	BRI	1.375	1.3	9	NSAIDs, DMARDs, GCs, Biologics	RIB	University
102	F	48	25	BRI	1.75	3.9	5.4	NSAIDs, DMARDs, Biologics	RIB/Homemaker	University
103	M	71	11	BRI	1.5	3.7	2.2	NSAIDs, GCs	Retired	University
104	M	78	1	BRI	0.5	4.7	0	DMARDs, GCs	Retired	College/apprenticeship
105	F	62	15	BRI	1.375	Inc.	Inc.	DMARDs, GCs, Biologics	Retired	University
106	F	62	2	BRI	0.75	Inc.	Inc.	DMARDs, GCs	Retired	School
107	F	37	9	NBT	1.375	3.5	3.6	NSAIDs, GCs, Biologics	Unemployed	College/apprenticeship
108	F	60	2	BRI	2.125	10	10	DMARDs	Paid work	School
109	F	33	3	BRI	2	1.6	5.8	NSAIDs, DMARDs, Biologics	Paid work	University
110	F	63	7	NBT	2.5	4.9	4.9	NSAIDs, DMARDs	Retired	School
111	M	74	7	NBT	0.125	1.8	5.2	DMARDs	Retired	School
112	F	48	23	NBT	2.625	4.6	7.6	NSAIDs, DMARDs	RIB	School
113	F	48	14	NBT	1	3.2	3.7	NSAIDs	RIB	College/apprenticeship
114	F	71	14	NBT	1.625	Inc.	Inc.	NSAIDs, DMARDs	Retired	School
115	M	45	2	NBT	1.25	2.8	6.7	DMARDs, Biologics	Paid work	School
116	F	55	27	NBT	1.25	6.5	7.7	DMARDs	Paid work	College/apprenticeship
Median and interquartile range (IQR) †=61 (48-67), ‡=10 (3-19), §=1.375 (1.125-1.875), ¥=3.7 (2.3-4.8) ¤=5.4 (3.7-7.7)										

Pt ID=Patient identification number; dis dur=disease duration; HAQ=Health assessment questionnaire 0-3 (3=most disabled); PtG=Disease activity score 0-10 (0=very well, 10=very badly); Pt pain=Pain assessment 0-100 (no pain-severe pain); NSAIDs=Non-steroidal anti-inflammatory drugs; DMARDs=Disease modifying anti-rheumatic drugs; GCs=Glucocorticoids; RIB=Receiving incapacity benefit; Inc.=Incomplete data

4.5.2 Thematic analysis

Analysis identified 219 codes which were grouped into six themes and smaller subthemes (Appendix E) that captured patients' experiences of RA stiffness. The four themes on the left are all interlinked, relating stiffness to RA, to behaviour and environment, as experienced both locally and widespread, and as highly variable. These themes influence and are influenced by the two themes on the right which capture a process of impact and management (Figure 4.1). Each theme and its subthemes are described below with patient quotes for illustration.

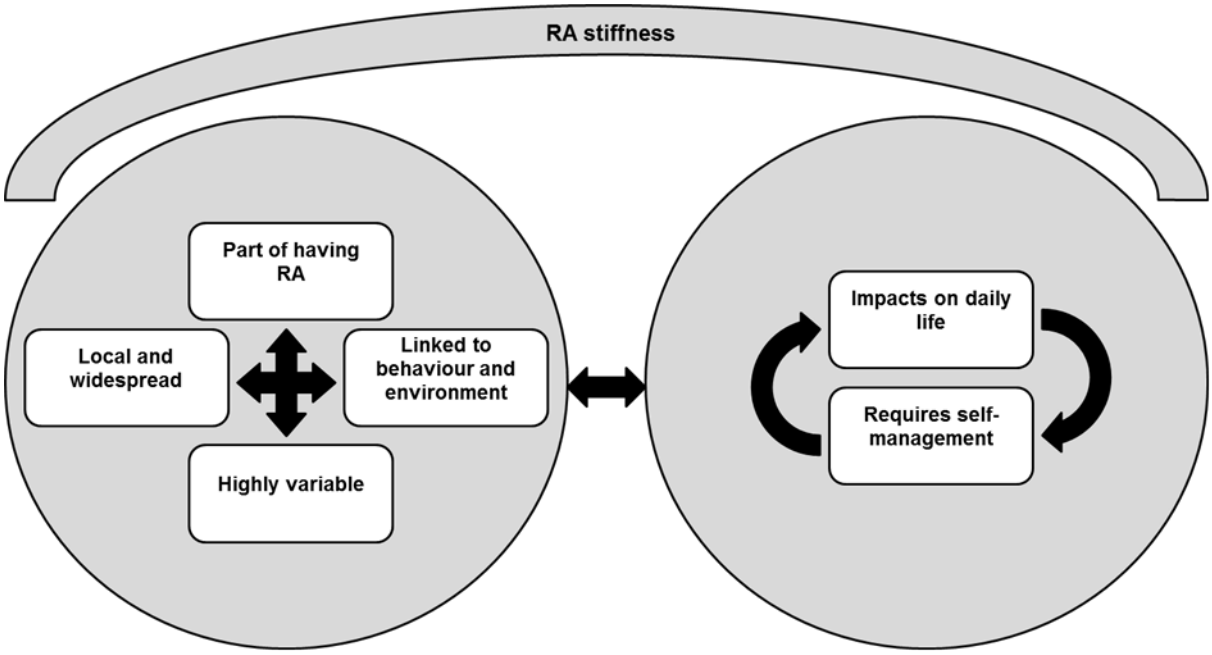


Figure 4.1: Conceptual diagram of the patient experience of stiffness

4.5.2.1 Theme 1: Part of having RA

Stiffness was discussed by patients within the context of RA. As such, discussions included a wide range of RA related topics that patients perceived as relating to stiffness.

4.5.2.1.1 Stiffness is a normal consequence of RA

Patients considered stiffness as part of their disease and a normal consequence of RA.

> *"All rheumatoid arthritis sufferers get used to a level of pain and a level of stiffness which they consider to be normal" [101]*

> *"[…] it's really one of the most obvious symptoms of the condition" [103]*

Some patients articulated stiffness being part of the broader context of RA using quite negative dialogue, indicating the negative influence of the disease.

> *"[…] it's all part of a picture that for me at the moment is just looking a bit black" [108]*

> *"And then I sit in bed telling myself that I do feel like death although I shouldn't use such a phrase […] Now I think that stiffness is part of that story" [103]*

There was discussion about the specificity of stiffness to RA. RA stiffness was considered different from stiffness as a result of exercise, due to differences in location and experience.

> *"[…] you swim a couple of miles and the next day […] I'd say I felt a bit stiff. That doesn't describe how I feel with arthritis in any way shape or form. It is a lot more uncomfortable" [108]*

> *"It's different to, you know if you did like a long run and you had sore muscles when you get up in the morning and you think 'oh God', and you've got that sort of stiffness of movement […] It's like you move like that but it's not in your muscles" [109]*

RA stiffness was also perceived as negative, unlike stiffness as a result of exercise.

> *"For me, muscle stiffness used to be, because obviously it's associated with working out […] it was like a nice 'ooh God yeah, I'm really stiff today' […] but with arthritis you know it's an on-going […] so it's a negative stiffness" [112]*

Although only considered by a small group of patients, the specificity of RA stiffness was also discussed in relation to other diseases. One patient described how she had a 'different' stiffness as a result of OA, differentiated by location.

> *"With me I am sure it is a different kind from my arms and my hands […] I think that is more the rheumatoid, but the stiffness I get from here [knees] and just sitting in any confined area, I think that is probably just general, because I have got osteoarthritis in my knees […] so that is a different stiffness" [106]*

Another patient however, stated that she was unable to differentiate between RA and OA.

> *"The thing is I've got rheumatoid and osteo, so I've got them both, so, I don't know which is which" [114]*

4.5.2.1.2 Stiffness varies with disease fluctuation

Patients identified that stiffness varied with fluctuations in disease activity and described this using a variety of terms. Many (n=10) used the term *"flare"*, and other terms such as *"not being controlled very well" [102], "bad period" [103], "bad patch" [113],* or *"bad time" [109]* were also used. During high disease activity or flare, patients described how all symptoms including stiffness were generally worse (*"it's much worse on a flare-up" [113]*). Stiffness worsening was discussed from a number of perspectives. Patients described how general physical functioning during periods of higher disease activity or flare was more restricted.

> *"Just everything I think when you get a flare and it's really bad. Again it's just your hands just don't work basically. They can't bend them, grip things, and obviously it's really painful and it makes everything awkward. When they are not so bad you can do basic stuff, you can pick up a kettle, you can do bits and bobs. There is a big difference between the two" [115]*

Patients highlighted that stiffness had an increased effect on activities of daily living during periods of higher disease activity or flare.

> *"[…] a frequent test, namely opening the car door […] is quite easy, not a problem. It causes a certain amount of discomfort but it's not difficult. But when I've been having a bad period […] then it's actually impossible, I'd get somebody else to do it" [103]*

> *"If I've got morning stiffness normally, it's just when I take a joint right to its end of range of movement […] in mid-range when you're kind of relaxed you don't feel any stiffness. It's just when you try to do something. When I'm severe it is as if I'm coated in a tight corset, even when I'm doing nothing at all. So it's like wearing, you know those Spandex underwear that Gok Wan always uses […] that's what it feels like without any of the benefits […] it is that sensation that you're, it's as if someone is just holding on to it and you've got to kind of fight against the movement to it" [109]*

Patients also perceived stiffness during higher disease activity or flare to be associated with other symptoms such as pain and inflammation.

> *"Oh I couldn't describe it to anybody how bad it was and that was stiff and painful as well" [106]*

> *"Yeah, I think it's because in flare there's always a lot more swelling with it as well, and for me swelling causes a lot more pain than just stiffness, especially in the bigger joints" [107]*

There was also an element of persistence with stiffness in periods of higher disease activity or flare. Here patients described how stiffness lasted longer, occurred at any time of day and for some patients there was no respite overnight which affected sleep.

> *"[…] in a bad time, I can't sleep because it's so bad overnight and then I'm woken every time I try to turn in bed […] So I struggle to sleep when I'm in a bad stage because the stiffness doesn't ease off overnight" [109]*

Some patients specifically highlighted that stiffness in periods of higher disease activity or flare could not be reduced with usual self-management techniques.

> *"So if I have a hot shower on a standard day I'm up and going and you know, that's pretty okay, whereas on a bad day I can't get it to reduce as well so it's just, it just lasts and I can't shift it" [109]*

In addition to amplified persistence, patients identified increased severity of stiffness in periods of higher disease activity or flare.

> *"I can tell if its, if I am going through a period when [I'm] not being controlled very well cause' the activity will increase […] I am stiffer either first thing in the morning or getting towards tea time in the evening" [102]*

The speed of onset of stiffness in times of higher disease activity or flare also featured in discussions.

> *"The other one, which is what happens when you are having a flare up which in my case happens, takes about an hour to 45 minutes and I know I am having a flare up and that is how quick it happens, I go from being mobile to being seized up very very quickly" [101]*

> *"Well I seem to be alright and then all of a sudden bang and I was that blooming stiff, to move or anything else" [105]*

The location of stiffness was another area of discussion with some patients feeling that stiffness was more widespread in the body during higher disease activity or flare.

> *"[…] it affects more joints than it does when I'm not so bad […]" [109]*

Other patients felt that stiffness was more specific in location during higher disease activity or flare.

"[…] for me the word that springs to mind is like acute, when it's like pain and stiff together […] it is so localised and you can't, you can't take your mind off of it because it's just so painful and stiff" [112]

Patients who also had stiffness from joint damage indicated that stiffness from mechanical and inflammatory processes felt different in terms of severity and persistence.

"I suppose a joint that's gone over, it's knackered, is a restrictive stiffness and pain but a joint that's flared is a completely different feeling [...] once you've got damage, you're always stiff" [112]

Within this, patients suggested that damage is a direct cause of stiffness.

"In that [damaged] joint you will always have stiffness and pain because you do. It doesn't completely go away" [112]

As a result of damage, patients were physically unable to move affected joints.

"[…] there's nothing I can do to move my hands. Nothing" [110]

Damage was identified by one patient as being the result of long disease duration and not having access to aggressive treatment from the start of the disease.

"The next generation coming through your system with rheumatoid arthritis are already being treated with these new drugs so they are not having the same issues with the mechanics of the body" [101]

Finally, it was suggested that damaged joints varied in severity in periods of higher disease activity or flare.

"[…] sometimes you can feel the start of it because it feels like somebody's dropped an Alka-Seltzer in your joint […] it is like a pins and needles, but it also throbs. There's like a pulse to it […] like a heartbeat in your joint, […] I say to him [consultant] I've got a fizzy joint" [112]

"I've had ops on my feet, I've had all the toes done, all the toes straightened. But my big toes on each foot, […] they are very stiff, very stiff. Um, I can't really wiggle that about […] And that's worse when I'm in a flare and when it's cold, because I don't think the circulation goes through to […] the end of your toes" [114]

4.5.2.1.3 Relationship between stiffness and other RA symptoms

Relationships with other RA symptoms were apparent, significantly between pain and stiffness. Some patients found it difficult to differentiate stiffness while others used both pain and stiffness descriptors (Section 4.5.3).

> *"[…] stiffness is the same as pain" [105]*

> *"[…] for me, stiffness is pain. I don't always have pain but I relate stiffness to pain" [112]*

Most patients however could discuss pain and stiffness independently and felt they were different yet related concepts.

> *"I mean the other thing is to sort of like pick out stiffness from all the other things that are part of the disease of fatigue and pain […] cause' they are intertwined" [102]*

> *"They're connected and related but they're not inter-dependent […] if I've got stiffness it's not guaranteed I've got pain" [107]*

> *"Stiffness can become pain [...] Initially, definitely they are separate" [108]*

The relationship between pain and stiffness was perceived to be stronger during higher disease activity or flare.

> *"I think they are separate but when, you know, when everything's sore, everything's swollen and everything's stiff, it's all kind of you know, in a bag together and then you're just in a pickle really" [109]*

There were less common discussions about the relationship between stiffness and fatigue and inflammation. One participant suggested that the relationship between fatigue and stiffness was stronger than between stiffness and pain as the physical sensation was similar.

> *"Um, I would say it was more [related] with fatigue rather than pain [...] its feeling just no energy and everything is just shutting down basically, seizing up, and just thinking I need to go to bed" [102]*

Other patients indicated that stiffness and fatigue were independent and distinguishable by timing and experience.

> *"[…] no I don't think stiffness and fatigue are linked at all, they're different. They come on at different times of day and they're really different in nature" [103]*

"It's different to if I'm fatigued and therefore my joints are achy and it's different to if I've had a busy day with the kids and I'm a bit sore" [109]

Some patients reported that inflammation related to and caused stiffness.

"[…] I just seem to have a permanent bit of swelling just under my knuckle, and that always creates a lot of stiffness" [107]

"Well that's stiff because it's swollen" [114]

However, inflammation was also described as independent and distinguishable from stiffness.

"[…] you can have the stiffness, i.e. the joint just feeling tight and heavy, whereas you can have inflammation which is obviously then causing it to be more difficult to move, because obviously there's more stuff in there" [107]

"In the same way, I had knee surgery before I had the rheumatoid arthritis, and that created a lot of swelling and your joint was stiff but it was visibly really big and swollen. It's the same sensation but without any visible swelling. Well I might have a bit but it's completely unrelated almost" [109]

A number of patients suggested that they may not recognise inflammation.

"I suspect it [stiffness] probably is [related to inflammation], but I haven't got any particular way of gauging that relationship" [103]

"I mean when you see somebody who knows something about it they say, 'oh yes, I can see the swelling there'. Oh, I didn't know there was swelling there. Oh I don't notice. It's only when they're quite fat that I would notice it you know, they might be painful, but I wouldn't really notice the swelling" [116]

4.5.2.1.4 Varying prominence of stiffness during the course of the disease

There were differences between patients regarding the prominence of stiffness during their disease. Stiffness was identified as being particularly significant in early disease and for some patients, it was one of the first noticeable symptoms.

"My rheumatoid arthritis started just after [son] was born [...] and it started with morning stiffness" [109]

"Yeah, I've had RA now for roughly about 18 months, came on rather quick, started on my feet, stiff, painful to walk. It was like that for a few months, seemed to go away, get a bit better. It just came back with, like a vengeance, started off getting stiff thumbs and first finger" [115]

Some patients reported that at disease onset, stiffness was particularly severe. It caused difficulty functioning and was perceived as relating to pain and inflammation.

> *"Well not so bad now, but […] when I had it first of all I could hardly move" [105]*

> *"Well when it first started […] one morning I woke up and I just couldn't move. I just couldn't literally move" [106]*

Other patients felt that during the early stages of their disease, stiffness was not as prominent as other symptoms.

> *"[…] when my disease was not very well controlled […] at the beginning, and very active, I didn't find stiffness was so much of a problem because it, the pain was sort of prevalent really and it went all the way through the day it was just continuous and all the way through the night so I didn't particularly notice stiffness as being particularly an issue" [102]*

> *"In the very beginning, I didn't have a lot of stiffness at all" [111]*

4.5.2.2 Theme 2: Local and widespread

Theme 2 captured the patient perception of stiffness being a physical bodily experience.

4.5.2.2.1 Affected body structures

Patients considered stiffness to relate to joints and many used the word joints to elaborate descriptions of the experience of stiffness.

> *"Yeah in, actually in, in the joint […] and right deep in the joint" [110]*

> *"[…] it feels like it's right inside […] well stiffness is right, it feels like it's right in your core like of whatever joint that is" [112]*

> *"I would say it was a joint and spreading out from the joint. So say that one there is quite stiff and I can feel it tense in there I suppose, tense in either side of the joint" [113]*

There was some ambiguity regarding the bodily structures (i.e. joints, muscles) related to stiffness. Some patients expressed uncertainty regarding stiffness being related to joints. One patient's description of stiffness in flare, which he likened to cramp, described muscles and tendons as being the relevant structures.

"You see it's the muscles and the tendons that are like contracting […] almost as if you are paralysed, you cannot move. So that's extreme stiffness, and I am not exaggerating that" [101]

However, interestingly other patients differentiated stiffness and pain by the body structure that was involved. This could be a result of the complex relationship between pain and stiffness.

"[…] stiffness I tend to relate more to the joints […] And pain tends to be […] it sort of radiates more out and it goes all up the tendons and ligaments and muscles" [102]

"I think it's more sort of muscly I think. I think well, the actual, the real bad pain is […] because it's swelled up, it's like all your tendons and stuff are kind of all pushed out and I think that's what causing the actual pain" [115] "And what bit is stiff?" [Halls] "Just, like, the joint" [115]

4.5.2.2.2 Location within body

Many patients described their experience of stiffness with reference to its location.

"Well I don't have a lot, but I have stiffness in my left hip and knee" [104]

"[…] it's nearly always in my hands and my shoulders" [107]

"It's mostly my hands and my feet […] I don't get the stiffness all over, I just get it very badly in my hands and my feet" [114]

It was suggested that over the course of the disease, the location of stiffness varied.

"[…] sometimes it is in my feet and sometimes it isn't, it's not in my feet at the moment so. It is a bit random, it does tend to move around […] I might be sort of six months with it really bad in my feet and my knees and then I might find that it is worse in my back and hips and then it might move up […] to my shoulders and my elbows" [102]

"Originally my knees were a bit stiff, but apart from that my knees have been fine […] Ankles are not much stiff, just [my] feet and it's like the front bit […] so the tops of your toes" [115]

For some patients, stiffness was described as being more of a whole body experience or more widespread, particularly during the morning or flare.

"[…] it's more general in the morning and more specified during the day. So it's more of an all-round stiffness because obviously you haven't moved around" [112]

"[…] stiffness when you're getting up, it feels like all up your arms and your legs and your whole body more" [116]

4.5.2.3 Theme 3: Linked to behaviour and environment

While patients related stiffness to their disease, they also associated it with their behaviour and environment.

4.5.2.3.1 Movement and stiffness

The topic of movement featured regularly throughout patient dialogue. Within this, stiffness was considered a result of both immobility and over-activity.

> *"[…] if I have had like a busy day, and I haven't been able to rest […] then I might find that it is creeping back in the evening as well" [102]*

> *"Oh it's always much more difficult to get up after sitting still" [103]*

In relation to over-activity, one patient suggested that stiffness was a signal to stop or change current behaviour.

> *"[…] it rings warning bells […] it tells you there's something wrong […] I mean your joints wouldn't be stiff if everything was fine. So it's, if the, alarm bells start ringing […] It's telling you like, you know, you know, you've done something wrong or you should be doing something differently" [110]*

Another patient suggested that after a restless night, her stiffness was worse in the morning which was not compatible with stiffness as a result of being immobile.

> *"Well, I am quite tired at the moment cause I am not sleeping very well […] And that, I tend to find that stiffness is worse in the morning but then I have probably been more active during the night, you know, so, you know really it should be better shouldn't it cause I haven't actually just sat and lay in one position for a long while" [102]*

Stiffness was also felt to be a result of a joint being in a fixed or restricted position.

> *"So I use the mouse, that's what makes my hand go stiff by holding the mouse" [113]*

4.5.2.3.2 Relationship between medications and stiffness

Many participants were keen to discuss their medications, within this there were discussions specifically regarding the relationship between medications and stiffness.

> *"[…] but since I've on the Humira injections I have found a difference […] I'm not as stiff as I was" [105]*

"[…] this morning it was about half an hour [...] and that's with taking the steroids, which does make it easier" [113]

Some participants reported dramatic effects of medications on their stiffness, in relation to enabling completion of activities of daily living and regaining normality.

"I was on originally which was Enbrel. Which was like a miracle drug for me […] I woke up in the morning […] and I was like 'oh my God', I could like move my hands, I got out of bed, like usually for me I have got to kind of like rock me-self up and then go like 'oh ah ah ah' like some old man. And I got up and it was just amazing. It was like every joint had been injected with some lubricant. I was like unbelievable and I just walked down the stairs one step after another like a normal person. It was like absolutely fantastic" [115]

Other patients indicated that medications had reduced stiffness to such an extent that it was no longer a problem.

"And now that I am on this infusion […] that seems to sort of alleviated that problem" [101]

"I have been on the Humira now for just coming up to 3 months [...] I feel better but I still suffer with the stiffness, especially in the mornings. Whereas on the Enbrel I never suffered any of that" [115]

One participant stressed that she felt the dramatic effect of medication on stiffness was ignored by the clinical team due to the lack of appropriate assessment.

"I kind of feel that it's sort of a lost entity because actually the drugs working, one of the things that they've really transformed has been my stiffness, but it's never been a measure that's kind of been considered [...] the one thing they've never asked me about is joint stiffness and the one thing I'm absolutely delighted about is that I can now get up and get him [son] up whereas I haven't for two and a half years because I can't do that in the morning [...] and like the nurses all know and that's great but if they measured it they'd be brilliant because I could then say 'Yeah, look', you know?" [109]

4.5.2.3.3 Lifestyle and environment and stiffness

A number of lifestyle and environmental factors were discussed as influencing the experience of stiffness. It was suggested that cold and wet conditions accentuated stiffness duration, severity and impact.

"Air conditioning is a killer by the way […] You get on a plane and they blow cold air at you, your joints will go stiff and painful within minutes" [101]

"[…] I do like the sunshine […] I just feel not so stiff everywhere, normally you know what I mean I'm quite good, whereas when it's tipping in rain I'm so blooming stiff I've got a job to move" [105]

A small group of patients suggested that certain foods and alcohol affected stiffness.

"I wouldn't eat cheese every day of the week because I found cheese erm makes me stiff as well" [105]

"[…] once I was at a party and I drank a bit too much and I noticed it in the morning, it was stiff but not like, not hangover wise, but […] I seemed to be stiffer" [115]

4.5.2.4 Theme 4: Highly variable

Theme 4 reflected the highly variable nature of the experience of stiffness. Patients emphasised this in relation to individual experience, time, duration and intensity; both within and between patients. Although there is inherent overlap in content with Theme 1 (Section 4.5.2.1.2), here these aspects capture the variability of stiffness from a broad perspective whereas Theme 1 captured specific aspects of change in stiffness relating to disease activity.

4.5.2.4.1 Stiffness is individual

The first feature of this theme was the very individual nature of stiffness. Participants felt that stiffness was a personal and subjective experience that meant they could only describe their own experience.

"[…] I've only obviously got my own personal experience" [107]

4.5.2.4.2 Temporal pattern of stiffness

There was considerable discussion regarding the timing of stiffness. Many patients suggested that stiffness was experienced in the morning.

"[…] stiffness, it's always there in the mornings, sometimes it's very bad" [103]

"Well first thing in the morning, I'm stiff as monkeys" [105]

"[…] it starts off when I get up […] That is the worse time of the day for me" [111]

However, for some patients stiffness was not related to the morning period.

"There is no particular time of the day and people will say well what about mornings? No" [101]

"[…] there used to be a question called morning stiffness […] it is a little bit more at the moment, but I never really associated it with mornings" [102]

The majority of patients highlighted a broader, variable temporal pattern with stiffness as lasting all day or recurring in the evenings.

"[…] I'd say I have the usual, stiffer in the morning and stiffer at the end of the day" [107]

"[…] stiffness is there give or take 24/7. It comes and goes in waves as it were, but at the same time, it never really goes away" [110]

"On a good day it is really just morning and evening" [109]

A small group of patients suggested that the temporal pattern of stiffness had changed over the course of their disease.

"[…] I used to have the morning stiffness only really. It's only really in the last few years that I've started getting evening stiffness as well, although the drug I'm on at the moment, that's now fading again so exciting!!" [107]

4.5.2.4.3 Duration of stiffness

There appeared to be wide variability in the duration of stiffness both between and within individuals.

"[…] sometimes is only 10 minutes and I can get rid of it really quite quickly and then other times it is just hanging on […] and I just gradually just shed it through the first hour or so of the day" [102]

During higher disease activity or flare, stiffness was perceived to increase in duration and persistence and was suggested to be more frequent in occurrence.

"[stiffness] will vary anything from about half an hour to, I have had up to about two/three hours, unless I've obviously had a bit of a flare up then obviously it can be all day thereabouts" [107]

"[…] if it's well-controlled then I probably get about maybe an hour, hour and a half of stiffness in the mornings […] if I'm not as well controlled, not necessarily in a flare, just not having a brilliant day, it's just 24/7. It just never quite shrugs off" [109]

Some patients were able to give a rough estimate of the duration of stiffness in the morning. These patients did however state that it was difficult to put a figure on duration.

"Erm oh it's difficult. A good hour. Yeah, a good hour. After a steroid it'll be shorter […] but in general it's about an hour" [112]

"Well there again it varies, um normally about two hours" [114]

Despite variation in stiffness duration for some patients, for another group of patients, stiffness duration appeared to be constant throughout the day.

"[…] there would probably always be some" [102]

"Yeah that doesn't alter for me […] From the minute I wake up to the minute I go to sleep" [108]

This led to suggestions that duration is not relevant when discussing stiffness. For some this was because of the unchanging duration, for others this was because of the impracticality of defining its duration.

"Ah yes, I think really the morning stiffness, which actually sort of comes on kind of during breakfast and lasts until mid-morning […] it's fairly fixed. So I don't think that how long it lasts is a very interesting variable […] But generally speaking it's always the same amount of time" [103]

"[…] and when you're asking specifically about stiffness in the morning, when does it start or when do you notice it from […] [Halls"] "[…] Time's not a factor because […] I might only sleep half an hour like, you know, or an hour at a time […] so the time of day isn't a factor" [110]

4.5.2.4.4 Severity of stiffness

Severity of stiffness was another area where the variability was apparent. The general experience of stiffness did not appear to be severe and one patient termed this *"mild stiffness" [101]*. Patients described how general stiffness could be managed and was not specifically related to pain or other symptoms, although it did still effect function.

"Yeah yeah, the mild stiffness is something you can overcome quite easily and you expect it to happen […] So if I sit in my car and drive too far without taking a break I know when I go to get out the car it's going to be a struggle for a couple of minutes. I am aware of it, I know it's happening, so I am prepared, alright" [101]

"[…] morning stiffness, once you get up and start moving it all tends to loosen up" [107]

Similarly to duration, during periods of flare or having less well controlled disease, the experience of stiffness was more severe, related to other symptoms and had a greater impact.

"Now, when you get what I call cramp stiffness, let's call it that then, that is pain like you have never experienced, it's like having an abscess toothache alright, it's severe alright" [101]

4.5.2.5 Theme 5: Impacts on daily life

Importantly, patients described and evaluated their stiffness in terms of its impact on a number of domains.

4.5.2.5.1 Daily life impact

Patients stressed the impact of stiffness on daily life. Activities of daily living such as eating and dressing were highlighted as being affected.

"[…] I'd end up eating a lot of soup because I just can't get my mouth open as wide to take even just a simple fork of food, and chewing just becomes a total non-starter. Yeah so I get joint pain in my jaw but also it's really stiff" [109]

"And it's like doing up your shoe laces and things like that you know […] And buttons, sometimes trying to do a button up, it's awful because you can't move properly you know. Well yeah I mean especially with this as well, you try and do a button up and this is stiff and swollen and it's difficult to do" [114]

For some patients, stiffness limited participation in leisure activities and hobbies.

"I'm making my step-daughter her prom jewellery at the moment […] whereas before I would've just made it in a night no problem at all but when I'm stiff [...] I can't do it because I can't pick up the bead or pick up the needle" [109]

Other patients described stiffness as disrupting normality.

"It's also true that the stiffness in the mornings means that I don't get around to doing things which other people do in the mornings, like going to the pool and swimming" [103]

"[…] my best time usually is late evening […] any time sort of past 7/8 o'clock-ish and sometimes before I go to bed I feel quite good. It seems to take that long to get back to some normality" [115]

Patients also explained that stiffness affected their work.

"I am not safe enough to be on a building site I don't think, I couldn't get up steps and stuff, things that I used to do" [102]

"[…] at work I struggle a lot more because I'm more desk-based on a bad day so I'm not getting up and talking to patients so much and then my hands will be okay

because I'll be typing and therefore they're just doing a little amount of movement whereas my lower half where I've been sat, just stationary" [109]

Interestingly, some patients perceived stiffness to have less of an impact following stopping work.

"[…] getting up to go to work was getting worse because it was taking longer to be able to get to move to be able to put the car in gear and, you know, things like that […] but, I think now because I'm not working […] it's easier, I can cope with it better" [113]

4.5.2.5.2 Physical impact

Physical function was considerably influenced by stiffness, including reduced mobility, balance, dexterity, grip and range and speed of movement. Patients described general inability or difficulty moving as a result of stiffness.

"Just, I mean a job to move really, your limbs and your joints, your fingers, erm you can move them but they just, I just find it sometimes initially quite hard to do" [106]

"[…] when I wake up in the morning if you could imagine, say you've got two bucket fulls of quick drying cement and you stick your hands in, and it's drying and you're trying to move your hands, you're kind of like forcing against it […] you can't sort of move very much" [115]

As well as patients highlighting general movement difficulty, they also reported restriction of quite specific movements. The first of these related to stiffness limiting range of movement.

"[…] but I would say stiffness is just about, just it's painful to move it and it is difficult to actually get any joint to full extension" [102]

"[…] this stiffness I think has only occurred in the last few months but it's not painful, and it's just that I can't exercise the full movement of my leg" [104]

Some participants explained how their dexterity was affected by stiffness.

"I mean it's like the other day I lost a screw out of my glasses and I could see this screw and it was down there, and do you think I could get my fingers to pick it up, I could not, I could not get my fingers to pick up this blooming stupid screw" [105]

Grip was another physical function that was limited by stiffness.

"You haven't quite got the right grip, you can't quite make a full fist like you would expect to make" [109]

"I suppose it makes things more difficult to do, like I wouldn't be able to open […] a bottle of milk say on a morning […] Because I can't grip" [113]

Patients explained how they were not able to complete tasks quickly.

"[…] sometimes I move and I'm not in pain, I'm just sort of slow getting going" [109]

"I think it's not being able to rush, I think or not being able to do something quick if you wanted to do it" [113]

Finally mobility and balance were also stressed as being affected by stiffness.

"[…] walking on stiff feet is just, it feels awkward, your balance is a bit skew, quite likely to fall over […] all of the bones in the feet aren't operating properly […]" [102]

"[…] in a chair, you can't stand straight up […] you tend to walk around with your hands on walls and […] you're looking for support all the time" [110]

4.5.2.5.3 Cognitive impact

A small group of patients described a cognitive element that appeared to be impacted by stiffness. Here patients described how thought processes were influenced by stiffness.

"[…] say for example, walking with a stick, you've got to remember to walk the right way […] if you go just on a kerb, because not every where's got dropped kerbs unfortunately […] you've got to remember to do it the right way" [110]

"[…] it sounds silly but I'd almost forgotten how to walk properly, because I was kind of always hobbling around […] because I'd gotten into the habit of doing it, […] I was thinking, I don't have to do that, I don't have to limp or hobble" [115]

4.5.2.5.4 Psychosocial impact

Stiffness was also felt to have psychosocial impact. One participant suggested that stiffness impacted on her personal image in terms of how she perceived herself and how she portrayed herself to others.

"[…] even my good leg I'm quite stiff and uncomfortable and look awkward, […] because also a lot of it is vanity that you don't want people to see you looking quite like that, do you know what I mean I'd quite like to kid myself and kid everybody else that I'm fit and I'm healthy" [108]

In relation to emotional wellbeing, many patients spoke about how the restrictions imposed by stiffness resulted in feelings of frustration.

"Disruptive, I can find it quite frustrating at times, especially when I'm up against I really want to get something done by a certain time or by a certain day of the week, or because I've got something else happening I need to get that done" [107]

"They [hands] wouldn't do anything. I mean, you could do a basic, you know you could pick up that pen, but you got no kind of dexterity or anything. It's just kind of just useless, which as I say I was really frustrated" [115]

For a small group of patients, stiffness appeared to cause low emotional states. Although this appeared to relate specifically to stiffness, it was also relevant within the broader context of RA, reinforcing the link between stiffness within the context of RA.

"So stiffness then to some extent is not so much a limiting, limited clinical term of experience, it could also relate to a wider feeling. Now looked at from another point of view, you could describe that as malaise, or you could describe it as hopelessness, or I've used the word aporia" [103]

"[…] this morning I put a couple of pancakes in and wanted some jam and couldn't undo the lid and didn't have any jam, now some days that'll make me cry, some days I think Jesus Christ I can't even have any jam on my toast, and that will finish me off, and other days like today I just think I can't have any jam […] that is borne more from frustration than pain and stiffness of not being able to do simple tasks feeling a bit useless" [108]

"[…] I got one thing wrong with me I don't want another thing wrong with me by getting depressed but it grinds you down so much that you end up becoming depressed as well. And it's like a real tough thing to deal with. You know, you've probably spoken to lots of people and they can, a lot just of get on with it but you can get on with it, but you know you can't, it just it makes your life stop […] you're virtually disabled, you can't do anything. And people don't understand it, just don't understand. They think 'Oh what, you've got stiff hands or something' they don't understand the pain and how it kind of affects you" [115]

Interestingly it was also suggested that emotional states impacted on the disease and subsequently stiffness.

"I think people's emotional state […] I mean I just ached everywhere and the stiffness I am sure it was worse then […] so I think that might have an impact" [106]

"[…] sometimes you think I can't do that so you just don't try. See what I mean? And then therefore makes the pain, stiffness, and it does because I, and a lot of it is up here [in your head]. You've got to try and be as positive as you possibly can because if you're not, well you might as well just give up […] I would say stress and anxiety. Being anxious about things. Definitely impacts on it" [112]

4.5.2.5.5 Pain impact

Patients also discussed how pain would result from performing movements restricted by stiffness.

> *"[…] I think for me it's a case of I can have the stiffness, I can be sat here and like my shoulder's feeling a little bit stiff, but I'm not having a lot of pain, whereas if I start moving it, if I tried to lift it now I'd be having pain" [107]*

> *"[…] the […] stiffness thing for me, my legs well you just end up hobbling, I hobble because the flexibility in your joints just, just isn't there and if you push it a bit too much to bend or use that knee as you would normally […] it just hurts" [108]*

Pain following stiffness restricted movement appeared to be exaggerated in periods of high disease activity or flare.

> *"[…] you must keep your range of movement, it's difficult when you are in pain to continuously move a joint into that position where it hurts. But then if you don't it gets stiffer" [102]*

4.5.2.6 Theme 6: Requires self-management

Patients articulated numerous strategies to self-manage stiffness. Strategies targeted a range of domains and were both direct (targeting stiffness) and indirect (targeting the consequences of stiffness).

4.5.2.6.1 Direct strategies

4.5.2.6.1.1 Movement

General movement, including walking and stretching, was highlighted by many patients as being an effective way to reduce stiffness.

> *"After you walk round for a few minutes and your joints start moving again, the stiffness goes" [101]*

> *"But as soon as you wake up […] it all starts up again and, and it is purely moving that's going to put it right" [110]*

Specifically, moving while still in bed was a strategy used to target morning stiffness.

> *"[…] before I get up out of bed I try to move all my joints […] to actually just try and get everything moving a bit" [102]*

> *"[…] I sort of sit on the edge of the bed and just move just gently before I get up […]" [106]*

General and specific exercises were also suggested as being beneficial to loosen up and relieve stiffness.

> *"[…] that would ease up if I actually exercised it, it's kind of loosening up all the fluids and everything, or that's what it feels like […]" [107]*

> *"I do a variety of different exercises […] So I've got a gyro ball, a very lightweight gyro ball […] I use it to just loosen out my joints […] so it kind of eases out the stiffness" [109]*

Supporting and physically manipulating joints was discussed.

> *"[…] you realise that you have worked into that sort of stiffness place […] and you need to actually lower it back into place and then just get it in a more comfortable position" [102]*

> *"Sometimes like this morning I had to physically bend my hands to get them to work because they just won't, they're kind of just locked" [115]*

4.5.2.6.1.2 Treatment or devices

Heat and cold techniques such as hot water bottles and ice packs were regularly employed by patients to directly relieve stiffness.

> *"Ice is really good […] It is good for comfort but then it's like a cold, hard stiffness then and pain. Whereas warm is, if I put a wheat bag on my ankle, it feels like ooh cwtch […] But then it doesn't get rid of the swelling, so you've still got the stiffness but it's duller, but the cold is like going back to that pointy and round as the cold reduces it but God that is really stiff then, as in the pain is more er, spikey" [112]*

For many patients, a hot shower first thing in the morning was a simple measure that enabled them to get going.

> *"First thing in the morning, to get me mobile […] hot water is wonderful [...] I can move then" [105]*

> *"So if I have a hot shower on a standard day I'm up and going […]" [109]*

Some patients stated how they used medications and painkillers as part of their morning routine to relieve stiffness.

> *"Well by the time I've done my hot water and had a shower and got me pills into me […] they start to work so, and you know it's nowhere near as stiff at 12 o'clock" [105]*

"Normally what happens is I get up, I go downstairs, have a cup of coffee, painkillers every morning, I've been taking those now for about two years." [114] "For your stiffness particularly?" [SH] "Yeah, yes […] and then it's about two hours after, then I start to loosen up" [114]

It was suggested by one patient that in a period of flare the only effective management was a steroid injection.

"So that's extreme stiffness, and I am not exaggerating that […] it seems the only way to resolve that one is to have a massive injection of steroid" [101]

A number of alternative therapies were also used to directly relieve stiffness. Again although these strategies were discussed in the context of stiffness, it was felt that these also related more broadly to RA.

"[…] she had some of these oils and that and she just went up and down my arms, 10 minutes and I could blooming move them […]" [105]

"[…] relaxation techniques, I have relaxation hypnotherapy CDs" [110]

The use of gadgets and aids also appeared to be an effective way to manage stiffness.

"And I've got a few aides to use, like a button thing and the zipper thing, oh and a brilliant thing to open a jar of milk, a carton" [112]

Some patients used splints to support joints, during a flare or as a reminder, although this strategy was not universally considered effective.

"I sometimes put splints on at night time because it makes it more comfortable, but then the stiffness is just brutal when I take them off" [109]

"[…] sometimes if it's just a bit bad in the morning, not too bad but a bit bad, I get up and I put my splints on and I'll keep it on for about an hour and a half, […] It's always like a reminder for yourself to be careful with your hands […]" [116]

4.5.2.6.2 Indirect strategies

4.5.2.6.2.1 Behaviour adjustment

Indirect strategies to manage the impact of stiffness included behaviour adjustment especially in relation to daily tasks and dressing. When the idea of working through stiffness was proposed to one patient she replied, *"I can work round it, really" [105].*

"People say oh that's a nice dress and I think yeah it's because I couldn't get my jeans on but you know, thanks […] And [son], I dress him differently […] like these are the trousers that my husband put him in this morning and I bought them as a spare pair because you can see they're fitted, so you have to kind of work them and you work the button and all that, whereas those are loose and you just pull them straight off […] and you just do it, you don't really think it I suppose. You just adapt in the morning" [109]

"Well yeah I mean especially with this as well, you try and do a button up and this is stiff and swollen and it's difficult to do. So I've got lots of things with no buttons on" [114]

Changes in behaviour also related to performing certain activities later in the day.

"Sometimes I don't have it [a shower] in the morning, because I can't do it. So I might have it in the evening and if I can't do it that day, I'll do it the next" [114]

There was discussion in the broader context of RA where patients aimed to ensure that RA did not take over their lives for example "*I can't let it beat me" [112].* One patient however, was changing career specifically as a result of stiffness yet this was not perceived negatively, rather as an opportunity.

"So I was thinking in some ways how it limits me and in other ways what it gives me opportunities to do […] I'm about to start a new career and in part that's because I struggle so much starting at 8.30 with my early morning stiffness […]" [109]

4.5.2.6.2.2 Prepare and plan

Patients also described having to prepare and plan tasks, including getting going earlier to compensate for slower movements.

"I had to go for an MRI scan […] and that was about 10 o'clock in the morning but it was the only one that they had and I thought well, I've just got to do it haven't I? I've just got to move myself a lot earlier" [105]

"But otherwise it doesn't interrupt my life too much, because I do try and pace as much as possible to try and make sure I can do as much as possible when I want to, or take prolonged rest periods before prolonged activity. And then rest afterwards again" [107]

4.5.2.6.2.3 External support

Patients discussed help from family and friends including support with particular tasks and encouragement or facilitation with the use of gadgets.

"[…] my daughter is brilliant […] she got me an electric jar opener that takes the lids off, because I couldn't undo them" [106]

"[…] and when I'm really stiff I have to wake my husband and he then turns me sort of by shifting me […] and I whinge away at him like nobody's business and then he puts the pillows back into the right place" [109]

For other patients, family and friends were flexible in the way that they planned or conducted their day-to-day life.

"[my children] who are 11 and 13 […] they just step into it and just, you know, do whatever, and one morning I'll offer them scrambled eggs or porridge and the next morning it's just cereal and help yourselves. And they just kind of roll with it" [109]

4.5.3 Stiffness descriptors and metaphors

4.5.3.1 Stiffness descriptors

Many of the stiffness descriptors used by patients were based around impact (such as movement difficulties and frustration) and these were discussed in Theme 5. Descriptors have been grouped under related headings (Table 4.3).

Table 4.3: Stiffness descriptors (number of transcripts identified in)

Descriptor heading	Descriptor
Unmoveable	Rigid (1) Seized (5) Encased (1) Locked (4) Frozen (2)
Restriction	Resistance (2) Restricted (3) Not fluid (3) Pulling (3) Not loose and limber (1) Need oiling (4) Constricted (2)
Discomfort	Sunburn (2) Pain (10) Throbby (3) Sore (3) Uncomfortable (5) Ache (6) Discomfort (3)
Tightness	Tight (5) Fat/full (3) Taught (1)

4.5.3.2 Metaphors

Descriptions of the physical sensation of stiffness were aided by the use of metaphors.

4.5.3.2.1 The physical sensation of stiffness in metaphors

Many metaphors were based around the difficulty of movement including unbalanced, heavy and lack of fluid movement, perhaps to articulate rigidity.

"I'm rather walking like somebody on stilts" [104]

"[…] it's like the treacle that was affecting the movement" [109]

"I'm walking like a robot first thing in the morning" [114]

Some patients likened the physical sensation of stiffness to feeling elderly.

"[…] you act as if you are like 80 and 90" [101]

"[…] hobbling about and well, being a bit like a little old lady really" [108]

"[…] it's like an old person feeling stiff in the morning" [112]

Other patients suggested that their joints required lubrication.

"And that's how my joints are. WD40 obvious!" [101]

"Like oiling a hinge […] Replace all the synovial fluid with WD40, lovely" [107]

4.5.3.2.2 Descriptions of stiffness using metaphors

The interview guide included a question asking patients to describe stiffness using the metaphor of an animal. This was not asked in every interview and not all patients responded to this question as some found it too unusual a concept to answer. However, it did appear to resonate for some patients and a number of interesting metaphors were described capturing a range of stiffness characteristics.

"I feel a bit like a tortoise sometimes [laughs] […] Slow and a bit precarious. I just feel sort of slow and a bit sort of um, I suppose a bit sort of to do with the feeling of being a bit encased within, sort of like being a bit, constricted, I suppose" [102]

"Any fossilised animal would probably be a good one […] fossils are very stiff, they don't move at all you see" [103]

"[…] obviously something like a lion […] because sometimes, you know, it roars at you […] to give you warnings, messages […] Stop, don't or do" [110]

"A hyena. […] Because they're a bloody nuisance […] They're nasty […] And a pain (Laughs)" [111]

"Well I would say stiffness is like a dog […] Because a dog is your companion so is constantly with you and […] he's constantly with you but you can control it […] So it's like, well it's a companion because it's always with you but then you do have some control because for me that's a dog isn't it? […] Always there, always with me, my companion but you've got some degree of control whether that be i.e. for me, drugs, attitude, do you know what I mean? So for me it's a dog" [112]

Other patients' answers appeared to describe the broader experience of RA.

"[…] it would probably be something quite nasty and aggressive in my opinion, because […] that's the way it affects me and makes me, personally the way I feel, it can make me feel angry and frustrated. So maybe kind of like a bear with a sore head sort of thing, maybe or a vicious cat or something […] Something you couldn't 100% trust sort of thing yeah, you don't want to stick your hand in the cage, because you know what's going to happen" [115]

Some patients used the opportunity to describe their experience using a metaphor to illustrate differences between stiffness as a result of different processes. Here some patients differentiated general daily stiffness from stiffness in higher disease activity or flare. One patient indicated that she would choose a different animal to describe stiffness in the two situations.

"Probably say something along the lines of a sloth […] but it doesn't quite cover it, if that makes sense […] they're quite slow and sluggish and that's kind of how stiffness makes me feel, especially when I've got it in lots of joints, it's just the slow and sluggish and heavy and, which is my impression. I mean sloths are slow and sluggish but I don't know about the heavy but it's just the impression of, or the mental relationship […] so the sloth doesn't quite cover it […] I'm not completely happy with my answer […] because I'd say during flare it should be worse, that sounds really disparaging of sloths doesn't it? [107]

Another patient described her stiffness in flare as an exaggeration of her usual experience based on the metaphor of a teenager.

"I think it's just an exaggeration of itself. I don't think it is a different beast […] But exactly like a teenage daughter who's lovely and then she's hormonal and she's still the same beast, just kind of a little bit more unpredictable, a bit worse, a bit more door-slammy and stompy, and then the next day everything's great" [109]

One patient elaborated on her description of daily stiffness as a dog and described different experiences of stiffness using metaphors and separated flare stiffness from morning stiffness and stiffness throughout the day.

> *"[…] acute flare up would be a right snappy little git. Yeah. He's be a real snappy one […] In the morning he would be like a cuddly er, like a cuddly, furry thing […] no he'd be a Pitbull if it was acute flare-up, like a proper you really don't know what to expect, a real aggression, could hurt you […] in the morning it would be a, not a Bichon because that's a bit specific but more of a furry overall and then throughout the day he'd be a Westie. He's lovely to stroke, you never know when he's going to have you" [112]*

Another participant described the difference between stiffness as a result of a mechanical rather than inflammatory process using metaphor.

> *"The inflammation around the joints has been taken away quite successfully […] Leaving joints which have been hammered for 20 odd years, which are absolutely wrecked, exposed. Now that's, the analogy I draw from that is like pulling up an old wreck [laughs] from the sea bed and as soon as it gets the air on it, umm it starts to decay" [101]*

> *"If you take a rusty old engine and try to move it, right, you will eventually move it if you apply enough pressure, that's stiff" [101]*

4.6 Discussion

Stiffness was reported to be a normal part of having RA, be experienced in joints and more widespread, be related to behavioural and environmental factors and have marked variability (including not being limited to early morning). It resulted in wide-ranging consequences, which had wide ranging impact on patients' daily lives and necessitated self-management.

The key finding was the emphasis that patients placed on the impact of stiffness. This finding is consistent with much of the literature relating stiffness to aspects of quality of life and work life (e.g. Phillips and Dow, 2012) (Section 1.5.2) and with other qualitative work that has focused on stiffness in an RA population. As detailed earlier (Section 1.6.2), Lineker *et al.* (1999) conducted a qualitative study that aimed to develop a patient centred definition of stiffness. The results from that study highlighted the importance of the impact of stiffness however, it resulted in no change to the way stiffness was conceptualised or assessed. In addition, other very recent work exploring the RA patient experience of stiffness identified the impact of stiffness on daily life as an overarching theme and broadly covered very similar domains to

this study (Orbai *et al,* 2014). This study was performed in the US and involved 20 RA patients across four focus groups. Importantly, one of the key messages of this article was the identification that patients shared a common language in their descriptions of stiffness impact (Orbai *et al,* 2014). This is relevant especially given that in all three qualitative studies (Lineker *et al,* 1999; Orbai *et al,* 2014; Halls *et al,* 2015) the patient experience of stiffness was identified as being individual and varied. Therefore, areas of unity and shared relevance are particularly significant when considering stiffness assessment and the importance that patients placed on stiffness impact has implications for PROM development. Taking this into account, the potential effectiveness of measuring stiffness based on concepts beyond severity and duration should be considered. Given that impact was how patients in this study defined and evaluated stiffness, and the identification of impact as a topic of shared relevance across patients (Orbai *et al,* 2014) there would be an argument for basing stiffness assessment on the concept of impact. This suggestion would also fit with the impact triad (Sanderson *et al,* 2011) which is a concept developed by patients and health professionals and recommends considering not only the severity of an outcome, but also its importance to patients and their ability to self-manage it. It is important to note that self-management is a term that has been defined in a number of ways and has no gold standard definition (Barlow *et al,* 2002). The theme relating to self-management in relation to the experience of stiffness captured a broad range of components including a variety of physical and psychosocial aspects. This fits well with Barlow's (2001) definition where self-management is defined as "[…] the individual's ability to manage the symptoms, treatment and psychosocial consequences and life style changes inherent in living with a chronic condition. Efficacious self-management encompasses ability to monitor one's condition and to effect the cognitive, behavioural and emotional responses necessary to maintain a satisfactory quality of life" (p.547). This definition will be used when referring to self-management throughout this research.

These data may also help explain reasons for the poor performance of current stiffness questions. As highlighted in Chapter 2, current stiffness assessment is generally based on items capturing the severity or duration of EMS/MS. However there is considerable variation in item wording, response options, format and time frame. For example MS duration items include various baselines ('from awakening' or 'from getting up') and endpoints ('initial improvement' or 'complete resolution') (e.g. Hazes, Hayton and Silman, 1993). The inconsistency of stiffness assessment is compounded by patient suggestion that stiffness does not solely relate to the

morning. This may also explain difficulties in trying to determine a start or end point and why it has been suggested that patients are often forced to report a cut-off time (Hazes *et al,* 1994). Furthermore, traditional simple questions appear to assume that patients are evaluating stiffness related to inflammatory processes. However, some patients in this study appeared to be able to identify differences between inflammatory and mechanical stiffness. Finally, in existing assessment there is no consideration of stiffness location, yet patients in this study reported stiffness in single and multiple joints, and also discussed a more widespread experience of overall stiffness. Therefore, in order to enhance stiffness assessment, individual items that capture patient relevant concepts such as impact, location, process and timing should be developed and tested for inclusion in an RA stiffness PROM. This is supported by results from other work (Orbai *et al,* 2014) and reinforced by discussion at OMERACT (Orbai *et al,* 2015).

Another interesting result from this study was the significance of stiffness at different times during the course of the disease, particularly in early disease, where for some patients, stiffness was particularly prominent. The relevance of symptoms in early disease has been an area of recent research partly due to better outcomes as a result of earlier treatment initiation (van der Linden *et al,* 2010). Recent work that produced a synthesis of qualitative literature focusing on symptoms in early RA, reported that stiffness was not regularly described in detail in the literature and also highlighted that there was no description of the meaning of stiffness or exploration of the concept of stiffness particularly at RA onset (Stack *et al,* 2013). However, stiffness was identified as a relevant symptom in early RA by some patients in Study 1. This finding reinforces the potential importance of stiffness in further work into the development of a 'symptom questionnaire' to identify individuals at risk of RA as early as possible, as suggested by Stack *et al.* (2013). It also identifies another area where effective stiffness assessment would be relevant.

Similarly to some of the research conducted in PMR (Hughes *et al,* 2012), and other work into the patient experience of stiffness (Rhind, Unsworth and Haslock, 1987; Lineker *et al,* 1999; Orbai *et al,* 2014), this work identified a relationship between stiffness and pain. Orbai *et al.* (2014) went as far as to suggest stiffness-pain interdependence based on their results, including that the experience of stiffness was overshadowed by pain, separation of the two symptoms is meaningless as they are so related, and as stiffness is less well understood, pain is a more useful outcome. This is an interesting argument especially from a measurement perspective, as there

is little value in assessing a concept that is already captured in other assessments. However, this argument fails to address the importance of stiffness as a relevant patient outcome that regardless of its potential placement in any hierarchy of symptoms is still commonly experienced by patients and regularly used clinically and in research. Although this study did indicate a clear relationship between the two symptoms, distinctions were identified including management strategies and medications that targeted pain and stiffness differently e.g. *"the heat will affect my stiffness for longer than it affects the pain" [109]*. It is also important to consider that certain patient-reported symptoms may be more relevant to some patients than others. Therefore, only by developing effective stiffness measures (that assesses the relevant concept/s) will we be able to adequately make decisions on the usefulness of stiffness as an outcome in RA. Further exploration into the usefulness of stiffness assessment in RA is supported in a recent paper that investigated which aspects of RA disease activity correlate best with PtG assessment (Ward, Guthrie and Alba, 2014). In the study, a full range of patient reported (pain, disability, MS duration, MS severity, fatigue, depression) and clinical (CRP, SJC, TJC) outcomes were evaluated. It was reported that MS severity was associated with changes in PtG assessment independent of changes in pain, indicating that pain cannot be used as a substitute assessment of stiffness. As a result the paper promoted further investigation into the use of stiffness severity as a useful RA outcome (Ward, Guthrie and Alba, 2015).

Finally, the investigation into metaphor provided a novel way of further investigating the patient experience of stiffness. The metaphors used elaborated these data, for example, some reinforced the stiffness descriptors relating to lack of movement and restriction (e.g. *"[…] walking like a robot […]" [114]*) while others highlighted the unpredictability of stiffness reflecting Theme 4 (e.g. *"[…] like a teenage daughter […] [109]*"). Although many metaphors were spontaneously generated in participant dialogue, some participants found describing stiffness using the metaphor of an animal difficult. This may have been due to the use of a specifically defined metaphor, although during a recent study that used colour to describe osteoarthritis pain it was reported that the task appeared easier for some participants than others (Wylde *et al,* 2013) indicating that it may be the concept of the metaphor that is difficult for some participants. Despite this, the use of metaphor provided useful data and a valuable insight into the patient experience of stiffness.

Although the study sample was small (n=16) with participants from two NHS trusts in the same city, it comprised a range of age, gender, treatment regimens and disease

duration. All participants were White British despite participants from other ethnicities being approached during recruitment. Although research suggests that stiffness is a relevant symptom in ethnically diverse populations (Kett *et al,* 2010), there may be cultural differences in the perception of stiffness, thus this is an area for further research. However, given the similarities between results in this study and that conducted by Orbai *et al.* (2014) where participants were ethnically heterogeneous, this may not be necessary. Slight differences in recruitment strategies across sites (Section 4.3.2.1) may have influenced the sample. Patients are more likely to participate in research if initially approached by their usual doctor (Newington and Metcalfe, 2014). However, as participants were recruited at two sites, any influence this may have had would have been small. Furthermore, this study was conducted in a transparent and systematic manner (Meyrick, 2006) with careful record keeping (Mays and Pope, 1995) to enhance rigor and reduce bias. Data saturation was achieved (Guest, Bunce and Johnson, 2006) and a key strength was the independent analysis by team members to enhance the validity of the findings (Mays and Pope, 1995; Cohen and Crabtree, 2008). Finally, following recommendations during the review process, the paper (Halls *et al,* 2015) was also reported in compliance with the consolidated criteria for reporting qualitative studies (COREQ) in an attempt to improve rigor and transparency (Tong, Sainsbury and Craig, 2007). The 32-item COREQ checklist was reported for this study and is available in Appendix F.

4.7 Conclusion

This study has demonstrated the relevance of stiffness to patients. Specifically, it included one patient who reported that the significant impact of her medications went unrecognised due to lack of an appropriate stiffness measure (Section 4.5.2.3.2). This important point now needs to be addressed in further research as currently, there is no clear evidence regarding the most appropriate measure to use to assess stiffness in RA and assessment relies on non-standardised items which do not appear to have been developed in accordance with recommended methods including collaboration with patients (Kirwan *et al,* 2007; Patrick *et al,* 2011a; Patrick *et al,* 2011b). Development and validation of a stiffness PROM would provide a standardised assessment method which could be implemented in clinical and research environments. It would open up the potential for stiffness to be included in the ACR disease activity core set (currently omitted omitted because it is not sensitive to change (Felson *et al,* 1993)) and would also address the OMERACT 2010 research agenda item to investigate stiffness assessment in relation to flare (Bingham *et al,*

2011). Most importantly, it will recognise and provide a systematic method of assessing a symptom that is important to patients and has a significant impact on patients' daily lives. As such, the next chapter will focus on the development of the content for a new RA stiffness PROM.

Chapter 5: Checking the conceptual model and investigating stiffness assessment (Study 2)

The previous chapter developed understanding of the patient experience of RA stiffness. This chapter focuses on the development of content for an RA stiffness PROM using focus groups to validate the findings from Study 1, and explore the patient perspective of RA stiffness assessment.

5.1 Background

Literature identified previously (Chapters 1 and 2; Study 1) demonstrated the relevance of stiffness to people with RA. However, current stiffness assessment tools are not standardised, poorly defined, have little measurement property evidence and do not appear to have been developed according to current standards (e.g. Patrick *et al,* 2011a) (Chapter 2). In order to recognise this relevant patient symptom and address the concerns associated with its assessment, the development and validation of a stiffness PROM is required.

Recommendations regarding the development of PROM highlight that collaboration with patients is essential for item generation (USDHHS FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b). Study 1 developed a conceptual model of the patient experience of stiffness in RA which will inform item generation. It is important to demonstrate rigor in qualitative results which can be achieved in different ways (Mays and Pope, 1995). Therefore, Study 2 provides the opportunity to test the conceptual model in a different sample of patients, using a different method of data collection. This will enhance the robustness of findings and may identify different views and opinions. Study 2 will also develop current understanding further by providing the opportunity to explore stiffness assessment with patients, which was not specifically investigated in Study 1. For example, impact was a concept identified in Study 1 as relevant to patients, and was suggested (Section 4.6) to fit with concepts in the literature that relate to measurement such as the impact triad (Sanderson *et al,* 2011). This study will enable investigation into the patient perspective into whether and how to incorporate this in stiffness assessment.

5.2 Aims and objectives

The overall aim of this study was to start working towards the development of the content for an RA stiffness PROM. The specific study objectives were:

- To test the results from Study 1 in different populations of patients using a different method of data collection to confirm and elaborate understanding
- To investigate the patient perspective of stiffness assessment
    - To investigate how patients would like stiffness captured in a questionnaire
    - To explore patient thoughts about how to improve current stiffness assessment
    - To explore patient thoughts about the impact triad as a possible conceptual basis for stiffness assessment

5.3 Methods

5.3.1 Focus group methodology

Focus groups are a method of collecting data on a particular topic from multiple individuals at the same time (Braun and Clarke, 2012). Researchers have suggested that focus groups have five characteristics: they involve people; specifically those with certain characteristics; they provide qualitative data; they involve focused discussion and they result in further understanding of the topic of interest (Krueger and Casey, 2009). Within the literature, a key use of focus groups has been to confirm not only the completeness of data gathered using other methods, but also to check that researchers have understood and interpreted data correctly (Finch and Lewis, 2003). Both of these uses were relevant outcomes for this study. In particular, it was important to test the interpretation of data collected in Study 1, and to generate discussion regarding stiffness assessment. For example, explore the link between the conceptual model of stiffness and the concept of the impact triad, and whether this was an acceptable and understandable basis for stiffness measurement. Additionally, in relation to investigating the patient perspective regarding stiffness assessment, focus groups have been suggested to be useful to elicit creative thinking (Lewis, 2003) through discussion, therefore providing an opportunity to capture ideas regarding patient relevant stiffness measurement.

5.3.1.1 Comparing interview and focus group methodology

Unlike interviews, which allow in-depth exploration of individual experiences, focus groups instead concentrate on the interaction between participants to explore the topic of interest (Finch and Lewis, 2003). The importance of interaction between participants, facilitated by the use of focus groups is a way of exploring participants' attitudes and the context in which they are set (Kitzinger, 1994). This allows

participants to influence, and be influenced by each other and reflects real life discussion (Krueger and Casey, 2009). This means that focus groups may provide a setting that is more natural than other qualitative methods such as interviews (Krueger and Casey, 2009) and it has been suggested that due to this, participants may communicate more naturally and use everyday language (Finch and Lewis, 2003). This would also provide the opportunity to further explore the language that patients' use when discussing stiffness, which would be relevant to capture in item development.

5.3.2 Participant identification and sampling

Participant identification and sampling was conducted as described in Chapter 4 (Section 4.3.2). The study specific PIS used during recruitment can be viewed in Appendix G. It is recommended that the number of focus groups should be based on consideration of the aims of the study and practical issues such as time, rather than quantitative estimates (Bloor *et al,* 2001). This is reinforced by Morgan (1997) who adds that the more heterogeneous the groups are, the more focus groups will be required. It has been suggested that the number of participants per group depends on the purpose of the focus group (Bloor *et al,* 2001; Barbour, 2008). Although in market research contexts, it is common to have as many as 12 participants per focus group, Barbour (2008) suggests that for social science purposes, eight participants' enables efficient moderation and analysis. At the lower end, a minimum of three or four participants is felt to be acceptable (Bloor *et al,* 2001). Barbour (2008) also acknowledges that practical considerations such as room size affect decisions regarding sample size. Given these recommendations, it was decided that the target for recruitment would be between two and four focus groups, each containing between four and eight participants. This target sample size was consistent with recommendations, suited the aims of the study and acknowledged practical considerations of time and the number of participants that could be comfortably accommodated in the rooms available. Additional focus groups could be conducted if required or data saturation was not achieved.

5.3.2.1 Site specific differences in participant identification and approach

As described in Chapter 4 (Section 4.3.2.1) there were slight differences in the identification and approach of potential participants at the two recruitment sites. Recruitment at NBT was consistent with earlier description but recruitment at the BRI differed slightly. As a result of the department being research active, multiple research studies often occur simultaneously. At the time of recruitment for this study there were

two other studies recruiting RA patients within the rheumatology department at the BRI. To minimise researcher presence in precious clinical space and to reduce burden on patients by being approached multiple times for different research studies, recruitment was conducted in conjunction with other researchers. Here, only one researcher was present at each clinic and recruited for all studies in an order dependent on the timeframe and inclusion criteria of each study. Study specific information was shared between researchers to ensure consistency.

5.3.3 Focus group topic guide development

A focus group topic guide was developed based on the Study 1 results, discussion with the supervisory team, and relevant literature. Topic guides were flexibly adhered to during each focus group. The original focus group topic guide (Appendix H) focused on the development of the content of the RA stiffness PROM and concentrated on two key areas 1) the experience of stiffness; and 2) measurement instrument development based on a critique of current questions used to assess stiffness. The original focus group topic guide was used in focus group 1 only and was then revised following feedback from the progression exam and discussion with the supervisory team. The progression exam is a University procedure to ensure adequate progress and enable discussion of the research with examiners external to the supervisory team. Feedback from the progression exam prompted further interrogation of the Study 1 analysis (Chapter 4 reports the updated analysis). It also resulted in changes to the focus group topic guide in an attempt to strengthen the study by ensuring that the topic guide matched the study aims. The revised focus group topic guide (Appendix I) also focused on the development of the content of the RA stiffness PROM and again concentrated on the same two key areas, but from a slightly different perspective. Initially it concentrated on confirmation and further elaboration of the Study 1 data and following this, it explored patient preferences for stiffness measurement instrument development. The approach to addressing the latter point was revised as it was felt that the original approach was too specific (discussed in Section 5.3.3.2). Given the changes to the focus group topic guide and to enhance transparency, the content of, and changes to, each section of the focus group topic guide and which focus groups these were relevant to, has been described below.

5.3.3.1 Focus group Part A: The experience of stiffness

In every focus group, participants were asked to start by discussing their experience of stiffness and the words they use to describe it. Generated words were written up on a flip-chart. Participants were then shown cards, each with a single stiffness

descriptor that had been used by participants in Study 1 (Table 4.3 and Section 4.5.3.2.1), or in the literature (Table 5.1). Words generated that were not already on cards were written on cards for discussion. Participants were asked to separate the cards into two piles; one including words that they liked and would use and another including words that they disliked and would not use.

Table 5.1: Stiffness descriptors from Study 1 and the literature

Descriptors from study 1	Descriptors from the literature
Rigid	Gelling (Wright and Johns, 1960;
Seized	Goddard *et al,* 1970)
Locked	Wooden (Helliwell, 1995)
Frozen	Grating (Helliwell, 1995)
Restricted	Stiff (Helliwell, 1995)
Not loose and limber	Weakness (Abramson, 1967; Helliwell,
Need oiling	1995)
Pain	
Ache	
Tight	
Heavy	

During discussion with the supervisory team, following the review of the Study 1 analysis (following the progression exam), a number of aspects were identified that would benefit from further clarification: 1) the specificity of stiffness; 2) the location of stiffness; and 3) changes in prominence of stiffness over disease duration (Table 5.2). As the focus groups presented an opportunity for further discussion of these questions from the patient perspective, participants were asked for their help in elaborating on these areas in addition to extending understanding of the patient experience. The discussion points in Table 5.2 were used as prompts in the focus group topic guide in focus groups 2 and 3.

5.3.3.2 Focus group Part B: Proposed ideas for PROM development

Participants were then asked about their thoughts, ideas and preferences in relation to the development of a stiffness measurement instrument. This was performed differently in focus group 1 and focus groups 2 and 3. In focus group 1, participants were asked to critique current stiffness questions (Box 5.1) and suggest how they could be improved. Questions were chosen from those identified in the literature review (Section 2.4.1) and represented both severity and duration. Questions were written up on a flip-chart and presented to participants for discussion.

Table 5.2: Areas of further discussion in focus groups 2 and 3

Area of discussion	Key discussion points
1) The specificity of stiffness	• Further exploration of the relationship to/independence from pain • Does this relationship change depending on disease activity? • Is stiffness a patient word?
2) The location of stiffness	• Does the idea of a wider/whole body experience of stiffness resonate? • Would this be an appropriate concept for assessment?
3) Changes in prominence of stiffness over disease duration	• The idea that stiffness was more prominent at certain points during the disease duration was proposed. Is this about changes in the severity/impact of stiffness or about changes in ability to manage?

In focus groups 2 and 3, participants were not presented with stiffness questions from the literature (Box 5.1), as during reflection and early analysis of focus group 1, this did not appear to provide the opportunity for creative thinking to capture ideas regarding the development of patient-centred items as had been hoped for. Instead, this appeared to limit and shape participants' thinking in relation to item development. Therefore rather than generating new ideas to improve the presented questions, much of the discussion focused on their inadequacy. Although this information was useful, it did not achieve the aims of the study. Therefore, in focus groups 2 and 3, participants were given an overview of the results from Study 1 and the link to the impact triad was proposed. This approach provided an explicit link between Studies 1 and 2, and was felt to be a broader way of generating discussion about developing a patient relevant stiffness measurement instrument. Each aspect of the triad was then discussed in relation to question development, focusing on the development of potential stem questions, response options, timeframe and layout. This enabled discussion of aspects relevant to patients but in a broader yet more targeted manner. Additionally, the aspect of severity within the impact triad is also a commonly used concept in traditional stiffness assessment, which linked the approach used in focus group 1, and focus groups 2 and 3.

5.3.4 Focus group procedure

All focus groups took place in non-clinical rooms in the Academic Rheumatology Unit at the BRI and were conducted by two researchers (Halls and ED). Each participant was greeted, introduced to other participants and provided with refreshments. All participants gave informed consent and completed a questionnaire pack as previously

described (Section 4.3.4.1). Once all participants were present, the researchers introduced themselves and explained the purpose and process of the session (Finch and Lewis, 2003). Some participants were known to the researchers through other research projects or departmental activities. Each group was asked if they had any questions before beginning. Each focus group was audio-recorded and transcribed verbatim. All recordings were transcribed by a transcription service and were checked for accuracy and anonymised by the researcher.

Box 5.1: Questions presented to participants in focus group 1

1. Were your joints stiff when you woke up today?

   Yes ☐ No ☐

   If yes, how long did this extra stiffness last?

   ☐ Less than 30 minutes
   ☐ 30 minutes to 1 hour
   ☐ 1 – 2 hours
   ☐ 2 – 4 hours
   ☐ More than 4 hours but less than all day
   ☐ All day

2. How would you describe the overall level of morning stiffness you have had from the time you wake up?

   No stiffness 0 1 2 3 4 5 6 7 8 9 10 Very severe stiffness

5.4 Analysis

Data were analysed using a deductive approach. Thematic analysis can be performed using an inductive or data driven approach (as in Study 1), or in a deductive or theoretical approach (Braun and Clarke, 2006). Deductive analysis is driven by theoretical influences such as existing theory, previous research, or coding frames (Braun and Clarke, 2006). To address the first objective of the study, to test the Study 1 findings, the coding frame identified in Study 1 (Appendix E) was used as a

theoretical framework for analysis. This approach enabled exploration of the extent to which these data were consistent with the original coding frame and where aspects could be elaborated to incorporate further understanding. The approach also attempted to enhance consistency across Study 1 and Study 2 analyses (Part A). To address the second objective, to investigate the patient perspective of stiffness assessment, existing literature was used to develop a framework based on considerations from the questionnaire design literature (e.g. Tourangeau, 1984; Fowler, 1995; Streiner and Norman, 2008). The framework included the broad categories of: stem questions and anchors, response options, timeframe, layout and format (Part B). Analysis was performed using the Nvivo 10 (QSR International Pty Ltd, 2012) software package and Microsoft Office Word 2013. Notes made during the focus groups were reviewed during analysis. Coding was performed by the researcher and was discussed with other supervisory team members and patient partners (GB and AE).

5.5 Results

5.5.1 Participants

Of the 143 individuals approached, 21 agreed to participate (15% recruitment rate). As is common with focus groups, five individuals did not attend on the day (Happell, 2007) therefore, 16 RA patients participated: 11 were female (69%), age range between 43 and 85 years and disease duration between one and 38 years (Table 5.3). Reasons given for declining participation were as described in Chapter 4 (Section 4.3.4) however, additionally multiple study recruitment will likely have affected participation.

Table 5.3: Participant demographic information

Pt ID	Gender	Age (yrs) †	Dis dur (yrs) ‡	Group	HAQ §	PtG ¥	Pt pain ¤	Current medication	Work status	Education
2101	F	71	1	1	0.75	0.2	0.2	DMARDs, GCs	Retired	University
2102	F	64	16	1	1.625	6.7	2.9	Analgesics, DMARDs, GCs	Retired	School
2103	F	43	2	1	0	0.3	0.1	DMARDs, GCs	Paid work	University
2104	F	67	4	1	1.875	8.7	9.3	NSAIDs, DMARDs, GCs, Bios	Retired	School
2105	M	85	6	1	1.375	4.7	9.1	DMARDs, GCs	Retired	University
2206	M	72	10	2	1.75	5.0	7.3	Analgesics, DMARDs	Retired	School
2207	F	59	3	2	1.875	3.4	6.2	Analgesics, DMARDs	RIB	University
2208	F	73	10	2	2.125	1.8	7.4	Changing	Retired	University
2209	F	78	8	2	0.75	7.2	3.8	DMARDs	Retired	College
2210	M	60	25	2	1.75	6.1	5.0	DMARDs, Bios	Retired/RIB	College
2311	M	72	2	3	1.875	5.2	5.3	DMARDs	Retired	School
2312	F	64	4	3	0	1.4	1.2	NSAIDs, DMARDs, GCs	Retired	University
2313	F	65	7	3	2.375	5.0	9.0	DMARDs, GCs	Retired	University
2314	F	55	5	3	1.625	2.4	1.2	DMARDs, GCs	Retired/student	College
2315	M	43	38	3	Inc.	5.3	4.8	No medication	Unemployed	School
2316	F	54	27	3	2.125	3.5	2.6	Analgesics, NSAIDs, Bios	RIB	College

Median and interquartile range (IQR) †=65 (57-72), ‡=7 (4-13), §=1.75 (0.75-1.875), ¥=4.9 (2.1-5.7) ¤=4.9 (1.9-7.4)

Pt ID=Patient identification number; dis dur=disease duration; HAQ=Health assessment questionnaire 0-3 (3=most disabled); PtG=Disease activity score 0-10 (0=very well, 10=very badly); Pt pain=Pain assessment 0-100 (no pain-severe pain); NSAIDs=Non-steroidal anti-inflammatory drugs; DMARDs=Disease modifying anti-rheumatic drugs; GCs=Glucocorticoids; Bios=Biologics; RIB=Receiving incapacity benefit; Inc.=Incomplete data

5.5.2 Focus groups

Three focus groups, each containing participants with a range of age, gender and disease duration were performed (Table 5.4). Focus groups lasted approximately 120 minutes each and were conducted with only the researchers and participants present.

Table 5.4: Comparison of participants across groups

		Focus group 1	Focus group 2	Focus group 3
Gender	**Male**	1	2	2
	Female	4	3	4
Age range (yrs)		44-85	59-78	43-72
Disease duration (yrs)		1-16	3-25	2-38

5.5.2.1 Focus group dynamics reflection

The interaction and engagement of participants differed between focus groups. The participants in focus group 1 were relaxed and easy to engage. They listened to each other and related what they had to say to others' experiences, which made it easy for the researchers to encourage relevant discussion. The participants in focus group 2 engaged really well together. All participants appeared to feel comfortable within the group very quickly and therefore were forthcoming with sharing information and experiences. This did mean that discussion strayed off topic on occasions but generally this was monitored by group members who brought discussion back on point. The researcher did not feel that the participants in focus group 3 engaged as well with each other in comparison to the previous two groups. Discussions were overly polite and participants tended to wait their turn to speak to the researchers rather than engage with each other. Due to this, discussions took longer and this was accentuated by the group being slightly larger than the previous two groups. In addition, the researchers found some members of the group difficult to keep on topic and unlike in other groups, this was not monitored by other group members.

5.5.2.2 Focus group Part A: The experience of stiffness

The six themes within the conceptual model identified in Study 1 were supported and reinforced by the focus group data. The conceptual model appeared consistent with the experiences of a new population of patients, using a different method of data collection which enhances the robustness of the findings. Themes and subthemes were reinforced by these data and an additional 23 codes were identified and added to the coding tree (Appendix J). To avoid repetition, only new codes are detailed below under each theme heading, with patient quotes for illustration. Key discussion points in the revised topic guide (Table 5.2) are also illustrated where relevant.

5.5.2.2.1 Theme 1: Part of having RA

Discussion relevant to Theme 1 related to the specificity of stiffness (Section 5.3.3.1). Here it was highlighted that stiffness was an acceptable word that was used by patients.

> *"Well, stiffness would be used, wouldn't it?" […] Everybody says it, don't they?" [2210] "Yes" [multiple responders] "Okay, so it's not the medical word?" [ED] "No" [multiple responders]*

It was also identified that stiffness was a usual part of aging which may detract from the specificity of stiffness in RA.

> *"[…] you get a bit stiff and kind of forget, we are all getting older and sometimes overdo it […]" [2103]*
>
> *"[…] I suspect a lot of older people get stiff anyway, you know what I mean?" [2312]*

As in Study 1, much discussion concentrated on the relationship between pain and stiffness and the differences and similarities between the two symptoms.

> *"And stiffness to me […] I define as, I can't bend something. You know? Like a stiff piece of paper, you can't bend it. But it is trying to separate out pain from stiffness, it's very difficult" [2102] "It is, it is difficult" [2104]*

It was suggested that difficulty describing stiffness influenced the difference between pain and stiffness.

> *"And you know when you said about the stiffness I thought well actually, we know what stiffness is but [laughs] is it this or is it that? […] It is difficult to define" [2104]*

It was suggested that the sensation of stiffness may linguistically fall under the heading of pain. This was also reflected in Study 1 where there was crossover between descriptive words for pain and stiffness.

> *"[…] if you're stiff, you know you're stiff because you've got a pain. It may not be the same as a sharp pain, but you feel it […] you're sitting down for a while and then you get up. It's stiff, it's awkward. But how do you know it's stiff? There's nothing saying, I'm stiff. You feel something. It's a form of pain, I think" [2105]*
>
> *"It's not a stabbing pain […] it's just there" [2208] "It's just there, and it feels tight" [2207]*

Consistent with Study 1, the relationship between stiffness and pain was more relevant in periods of high disease activity or flare. This was described by one group as *"painful stiffness" [2210]*.

> *"Because at the moment I'm in a really bad flare up […] I've got a lot of pain and a lot of stiffness […] But trying to define which is which, I don't know" [2104]*

Furthermore the specificity of stiffness in RA was challenged by one participant who suggested that pain may differ across conditions, but stiffness was similar.

> *"The stiffness, yes, yes, yes, yes, very similar, but the pain is different […] my pain for rheumatoid is in the joints, whereas […] with the polymyalgia it's in the muscles and so it feels different, if that makes sense" [2208]*

Despite the close relationship between pain and stiffness, participants in focus group 1 suggested that it was difficult to respond to questions where stiffness and pain were asked within the same question.

> *"I find it very difficult when you are asked 'are you experiencing pain and stiffness?' Because I don't know what it is they are asking […] we've all got a different response and we do know there's a difference" [2101] "Are you often asked it together, then? In one sentence? [ED] "Oh yeah. Invariably. Related to osteo and RA things […] Always, pain and stiffness" [2101]*

Furthermore, it was highlighted that impact based tasks were more affected by stiffness than pain.

> *"If we're going to separate, it's the stiffness that's preventing you being as, you know doing that" [2104] […] "That was a fastening the bra movement for the tape" [2103] [laughter] "I'm not going to not put my bra on because it's going to hurt me. I'm going to find it difficult, because I'm too stiff" [2101] [agreement]*
>
> *"I think the stiffness is more of a factor than the pain, in getting dressed or undressed […] I find it impossible to do up my top button on a shirt if I've got to go somewhere formally. I can put the tie on, I can do all that. But I cannot do that button. Just can't. That's it" [2105]*
>
> *"Sometimes stiffness is prevention of doing things, like I used to enjoy sewing and threading a needle. I can't do that any longer, and that's the stiffness, it's not the pain. There's no pain in threading a needle, but I just can't do it because my fingers, parts of me just don't work in the way that they should do […]" [2208]*

The varying prominence of stiffness was also briefly addressed (Section 5.3.3.1). Consistent with Study 1, some participants felt that stiffness was more prominent at different times throughout their disease duration. In focus group 1, one participant

suggested that her stiffness had been more prominent in early disease but this had changed over the disease duration.

> *"[…] I don't have that sort of stiffness. I mean I used to, but when I was put on medication, the stiffness totally went for years and years and years […] it only happened in the morning for over an hour […] and after that, I was totally normal" [2102]*

Another participant suggested that in early disease other symptoms were more prominent than stiffness.

> *"[…] when it first came on I think I had stiffness but the pain was the overriding factor so I don't think I really recognised that I had stiffness and also it was all new to me […] now I can get stiffness and it could be just stiffness or it might be stiffness and pain" [2316]*

This participant went on to suggest that the ability to manage stiffness may be more effective in individuals with longer disease durations.

> *"[…] I think those of us who have had RA for a long time, we manage the stiffness" [2316] "That's right" [2311]*

Overall, the identification of stiffness as a usual part of aging, the association of stiffness as a symptom in other diseases, and the complicated relationship between stiffness and pain all reinforced that the focus on stiffness should be emphasised throughout any newly developed tool. Although it was highlighted that impact based tasks may be a patient-relevant way of separating stiffness and pain. The cause of change in the prominence of stiffness was also briefly discussed.

5.5.2.2.2 Theme 2: Local and widespread

Discussion relevant to Theme 2 related to the location of stiffness (Section 5.3.3.1). The idea of stiffness affecting the "whole body" was discussed.

> *"I used to get so embarrassed because my whole body was stiff" [2207]*

This was suggested to occur during periods of high disease activity or flare.

> *"It was a flare up […] I never had no swelling" [2210] "Okay, but you did have stiffness?" [ED] "Oh yeah! It stiffens your whole body up" [2210]*

This expanded the idea from Study 1 that stiffness in many joints may indicate severity, highlighting that stiffness in the "whole body" may also be relevant, particularly during high disease activity. This also identified patient-centred wording for inclusion in item development.

5.5.2.2.3 Theme 3: Linked to behaviour and environment

Within Theme 3, participants reinforced that environmental factors influenced the experience of stiffness.

> *"I think the dampness does something though isn't it" [2311] "[…] there is definitely evidence that it's something to do with the pressure" [2314]*

Although, the influence of environmental factors was not relevant for everyone.

> *"[…] mine also I don't think it's affected by the weather" [2312]*

5.5.2.2.4 Theme 4: Highly variable

The key additional discussion point highlighted within Theme 4 was the relevance of stiffness during the night.

> *"Sorry, but I find stiffness during the night. Never mind when I wake up" [2105]*

It was suggested that stiffness during the night could affect sleep especially in relation to movement during that time.

> *"I noticed at the very beginning I think you said stiffness wakes you at night […] Is that something that would be important?" [ED] [agreement] "I think you can twist and turn all night because of it" [2206] […] "I think I wake up about every 2 hours because I'm stiff, and […] if the stiffness wakes me up I can't turn over in my sleep" [2208]*

5.5.2.2.5 Theme 5: Impacts on daily life

As in Study 1, there was lots of discussion about the impact of stiffness. One focus group highlighted that it was not only dressing but undressing that was affected. This may be relevant to include as patient-centred wording during item development.

> *"You don't include there getting undressed […] I can get a tight sleeved thing on [...] but it's going to be more difficult to take it off" [2101]*

Another focus group suggested that stiffness makes daily life activities more of an effort. This related to discussion with one patient partner (AE) during a team meeting while reviewing the first two interview transcripts from Study 1 where it was discussed that stiffness required effort while pain did not.

> *"[…] but stiffness is an important part of it because you may not feel like putting on your glad rags […] you know, you're so stiff it's difficult […] I just don't want to change and put something fresh on" [2208] "you just can't be bothered like, you know" [2206] [agreement]*

> *"[…] as the pain is wearing down, then you find that its more and more difficult to move, it's more difficult to get up [and] my joints are more slow when I'm walking, so then walking becomes more 'Oh I just want to get somewhere' but I'm having to do just very small steps because I can't go very fast" [AE] "So is there an issue there with effort?" [SH] "Yeah" [AE] "The pain doesn't require extra effort" [SH] "But the stiffness does I think definitely" [AE]*

Loss of strength was also discussed in all focus groups in relation to physical impact.

> *"Some of it is just total loss of strength, isn't it?" [2104] [agreement]*

> *"So is stiffness sometimes related to strength then as well?" [ED] "Yes" [Multi] […] "Once you're stiff you lose all your strength" [2210] "It seems to sap you" [2208]*

Although there was some debate about whether loss of strength was directly a result of stiffness.

> *"One word that you mentioned 2312 that I don't know if it's to do with stiffness, was you said strength […] Is feeling a lack of strength part of stiffness?" [ED] "It can be but also […] we lose strength anyway" [2316] "I don't know about that" [2312] "You lose the muscle don't you […] [2315] "But there is also, so we lose strength quite quickly as you say and the strength is also affected if you're in a lot of pain with your hands for example, you haven't got the same strength but when they're stiff, you know, you can't move them as easily so then you've lost a bit more strength" [2316] "Like a chicken and an egg situation" [2314]*

There was also elaboration regarding the impact of stiffness on cognitive processes. Participants suggested that when you are stiff, your body does not move as expected.

> *"Bits of my body won't move when my brain is expecting them to move" [2101] "That's fair enough" [2105] "So you try and nothing happens or? [2102] "It's the stage before I try, it's an awareness that it's not working" [2101] [laughter] […] "And it's not fear that it's going to hurt. You just suddenly think, oh […] for me, like the first time I couldn't get out of the bath which was a few years ago. Almost as if I'd forgotten how to do something very regular" [2101]*

This was furthered by suggestion that the automatic instinct of movement was impaired by stiffness and movements required extra thought to complete.

> *"[…] it has been that I couldn't get up from the chair without thinking about it" [2101]*
>
> *"You've suddenly got to think, literally think, differently […] Because the stiffness isn't giving me that extra couple of inches […]" [2101] "I find stairs particularly difficult, whereas automatically you'd raise your foot enough to clear the step […] I've got to think about placing my foot […] I've got to consciously think, I've got to lift my foot. I shouldn't have to do that. I mean that's, it's instinctive" [2105]*
>
> *"It's almost like the messages aren't getting from the brain through to where they should be going, if that makes sense" [2208]*

Finally there was discussion about the psychological impact of stiffness. This was particularly discussed by one group where one quite newly diagnosed participant highlighted that stiffness caused worry.

> *"I'd say sort of looking at these frozen and seized, that's certain. But there's no, you know, I'd say emotional, anxiety, I've always been very scared. You know you'd be panicking, what's this leading to? [2103]*

This discussion continued with the suggestion that having guidelines about expected amounts of stiffness would be useful to reduce worry.

> *"So it's something I'm very aware of, so the slight stiffness when I wake up and I have mentioned it a couple of times when I come for review and they [clinical team] just say 'look if it's only five, ten minutes, you're alright' […] So it's nice to have some, if you get sort of look, 30 minutes, then start worrying. I find that really helpful […] because otherwise, you just do worry" [2103]*

However, worry was not something that was relevant for all participants, although this was discussed in the broader context of RA rather than being stiffness specific.

> *"One of the other things I think that makes a difference is, I am, um, unlike you, I've never been anxious about the RA and how it goes" [2102] "What's your secret?" [2103] [laughter]*

One participant also suggested that stiffness caused embarrassment which although only mentioned by one participant was a powerful statement.

> *"But when I got up, [laughs] I used to get so embarrassed because my whole body was stiff, and okay, people get stiff when they've been sat down, but I had to sort of, a stop or so before, I had to kind of start getting myself, you know, ready [laughs], and then, oh right, here we go, yes, next stop, and then push up, but*

nothing was working. Push up and stand up, and then press the bell, and [laughs] the walk down the bus to get off was so embarrassing, because I was so stiff, and I thought well, there's something wrong here, I shouldn't be like this" [2208]

5.5.2.2.6 Theme 6: Requires self-management

Similar points were discussed in Studies 1 and 2. However, during the focus groups there was further discussion about whether movement always reduced stiffness. This may reinforce the suggestion from Theme 1 that stiffness in periods of higher disease activity or flare cannot be reduced with usual self-management strategies.

"So if your knees are stiff, even just moving them a little does it help to ease the stiffness gradually" [2316] "Erm sometimes I can have stiffness that literally just lasts a few minutes or it might last a while […] but whether movement helps I don't really know because with my hands, they were particularly bad over Christmas and there was inflammation as well and it was painful […] but I mean I use my hands, you have to use your hands so it wasn't as if I wasn't using them and using them wasn't helping at all" [2314]

5.5.2.2.7 Stiffness descriptors

Participants used a variety of words and phrases to describe the experience of stiffness which varied across groups (Table 5.5).

Table 5.5: Stiffness descriptors generated by each focus group

Focus group 1	Focus group 2	Focus group 3
Fizzing	Embarrassing	Can't move
Anxiety	Tin man	Tense-up
Set in stone	Locks	Frustration
Loss of strength	Painful stiffness	Moving
Frustrating	Numbness	Mechanical
	Creaking	Seized up
	Clicking	

In addition to the independently generated stiffness descriptors, participants were asked to discuss descriptors used in Study 1 and from the literature. Table 5.6 indicates patient preferences for the descriptors. Here *restricted*, *need oiling*, and *weakness* were liked across all focus groups however, none of these words were used independently by participants. There was general dislike across the focus groups for *gelling*. Apart from the word *stiff* or *stiffness*, there was no single word or phrase that was used consistently across focus groups. As highlighted above (Section 5.5.2.2.1), stiffness was regarded as a patient relevant word rather than a medical term.

Table 5.6: Patient preferences for stiffness descriptors

Descriptors	Focus group 1	Focus group 2	Focus group 3
Rigid	×		×
Seized	✓	✓	?
Locked		✓	×
Frozen	✓		×
Restricted	✓	✓	✓
Not loose and limber	?	×	?
Need oiling	✓	✓	✓
Pain	✓	✓	×
Ache	✓	✓	×
Tight			✓
Weakness	✓	✓	✓
Heavy		?	×
Gelling	×	×	×
Wooden		✓	✓
Grating		✓	×
Stiff	✓	✓	✓

Liked descriptors=✓; Disliked descriptors=×; Mixed opinions=?

5.5.2.3 Focus group Part B: Proposed ideas for PROM development

Each heading of the deductive analysis framework was populated by participant responses and a coding tree was developed (Appendix K). Each heading is described below with patient quotes for illustration.

5.5.2.3.1 Stem questions and anchors

5.5.2.3.1.1 Relevant to the individual

The individuality of stiffness, initially highlighted in Study 1, was emphasised by focus group participants who felt that measurement should reflect this.

> *"It's how you feel, not the average or somebody else" [2105]*

> *"It's all down to the individual. I shouldn't think there's two people the same" [2104] [agreement]*

It was suggested that wording questions to reflect the individual would be relevant, particularly in the context of whether people's current experience was typical for them.

> *"I did [a] questionnaire for one of your colleagues the other day […] I was like, last week I'd had a brilliant week. The two and half months before that had been an absolute nightmare and I wanted a box to say, is this usual?" [2103]*

However, it was suggested by one group that wording questions on a personal level may be difficult given the variability of the experience.

"I don't think you could say usual, but if you are on direct access and you want to come in and see somebody and then by the time you turn up it's all [fine] you know […] You say, I came in because […] everything was on fire last week" [2103]

In one focus group it was suggested that measurement may need to consider changing perspectives over time or response shift. One patient discussed how perceptions may be different for newly diagnosed patients and those who have had RA for longer duration.

"I think with you though, you're long established rheumatoid […] With us we're new diagnosed […] I think it changes doesn't it" [2311]

Another patient mentioned discussing response shift during a clinical appointment in relation to fatigue. She suggested that she based her assessment on how she felt before she had RA, while others may assess their symptoms based on a readjusted perception of what is usual for them now they have RA.

"[…] they ask you all these questions and because I'm still relatively newly diagnosed I couldn't understand […] am I supposed to be comparing it with how I feel on a good day or am I supposed to be comparing it to how I used to feel? And he [consultant] talked about something called a response shift whereby you get used to something and that becomes your norm […] I still know how I used to feel so for me I still always measure it compared to how I used to be because that is my measure […] But for some other people they've got used to being that this is my normal" [2314]

5.5.2.3.1.2 Impact

Impact was specifically discussed in relation to stiffness assessment. It was suggested that to capture stiffness, questions could be worded around its impact on daily tasks or movements.

"I know it's straight forward questions but it's serious questions for people that can't do it […] Comb your hair, brush your teeth, general daily, you think of what you do every day when you get up" [2210]

This was furthered by consideration of the increased time required to complete certain tasks or activities.

"I think the time element as well, how long it takes you to do these things" [2207] [agreement] "That's another thing, even though you may be able to do these jobs, does it take you, how much longer does it take you to carry these out?" [2210] "Okay. So say something like, my daily tasks take me longer than usual?" [ED] "Definitely" [2210] [agreement]

Another participant suggested that quality of life impact was a useful assessment for any symptom.

> *"[…] no matter whether you're discussing stiffness, pain, at the end of the day for me personally it boils down to my quality of life within the last week […] so what is your quality of life or how has it affected you within the last week?" [2313]*

Despite the importance of impact being captured in measurement being stressed by participants, a number of considerations were also identified. It was noted that some impact based questions were gender specific.

> *"And for ladies, fastening your bra" [2101] "We would have to have gender questions then, wouldn't we?" [ED] [laughter]*

It was also suggested that some impact based questions were time and location specific.

> *"But I think that […] getting dressed is probably in the morning. So if you're stiff in the morning, that's going to be really difficult for you. So I think that's perhaps not representative of later on in the day. That is morning specific question, isn't it, really?" [2103]*
>
> *"It depends, doesn't it? I mean, if you're not going to use those joints in the things you want to do, you're not going to have the problem, perhaps. The converse could apply" [2105]*

Another participant noted that impact questions should capture the degree of difficulty rather than whether or not tasks and activities can be completed.

> *"And I was still at work then and I never missed a day. So despite being stiff for an hour, I was still able to shower, get myself dressed and get to work in time. So, although I was obviously aware of it, it did not actually impinge on my life" [2102] "But it wasn't easy I bet" [2103] "The difficulty was still there" [2105] "Yes, but you just, you do it" [2102]*

The concept of the impact triad was proposed to participants in focus groups 2 and 3. In general this was well received by participants and it was felt that it would be an acceptable concept on which to base measurement.

> *"[…] if you don't want lots of questions […] is there a way of condensing it all down? [Halls] […] "Yes, I think you've got the answer there in front of you now […] Four questions there, there's four answers there, so how does it affect your self-management, how does it affect you severely, how important is it and what is the impact on your life?" [2210]*

"Yeah I would agree […] because it can vary, it really does, the importance is the impact that it's having not necessarily the severity, it's the impact that is the all-important" [2314]

It was felt that the impact triad could relate to any symptom and could reflect the individual experience.

"Any symptom, any activity" [2316]

"I think each person is different to everyone else aren't they" [2311] […] "I think that's possibly why that might work isn't it" [2316] […] "I was just going to say that, it wouldn't really matter that our experiences are all different because the important thing is the impact it's having on us as an individual" [2314]

Specifically, it was suggested that importance could be graded.

"Sorry, the importance, I think you could grade that as well […] Not very important, and then maybe you could put something else in, it hardly bothers me because I live with it, of medium importance maybe okay, it's there, but I get on with it. And then maybe ending up with another two – actually, it is really important because I can't do my daily functions, you know, your activities of daily, I can't do it, because I am so stiff" [2207]

5.5.2.3.1.3 Stiffness after a period of immobility

Focus group 1 participants suggested that it would be relevant to ask about stiffness following a period of immobility.

"[…] 2104 said earlier on that you know if you're sitting watching a film on the telly for a couple of hours, [when getting up] oh you're a bit stiff [laughs]" [2102] "That's right, it's like getting up again" [2105] "So it's not only during the night or first thing in the morning, it's also" [2102] [agreement] "It could be anytime" [2104] "Yes, you're right. Sitting here, for example" [2105] [laughter] "Exactly!! After a period of immobility, whether it's asleep or you're awake" [2101] "That's a good word, I like that […] Immobility. If you've been immobile for, I don't know, an hour. Whatever. Certainly longer. Then, how are your joints then? Nobody has asked that" [2105]

This was reflected in participant experiences in focus group 3 where stiffness would result from sitting for a period of time.

"[…] standing up after you've been sitting for a little time I have to stand up for a, you know, I say to friends "I'm just kind of just having my moment standing up" when I have to gather, my legs have to kind of gather themselves ready to move. I couldn't just always stand up and […] move straightaway" [2316] [agreement]

5.5.2.3.1.4 Timing and temporal pattern

The timing and temporal pattern of stiffness was reinforced as relevant for patients. However, current questions regarding the duration of morning stiffness were identified as being difficult to answer, inaccurate and hard to remember or quantify.

> *"I don't know how long […] you can't put a time on it" [2105]*

> *"[…] I have been asked how long does the stiffness last" [2209] "Yes. Is that an easy question to answer?" [ED] "No, it's not, because quite often by then I've forgotten how long it lasts" [2209]*

> *"Well, I don't know, this is when you come to the doctors and they say how long does it last, well, it's about that long but it's a guess really" [2209] "Yes and you suddenly realise you haven't got it then" [2208] [agreement]*

Patients suggested that they were unsure what the start or endpoint for these questions were.

> *"Do you take notice though of how long it takes you to get rid of your stiffness? Because I don't" [2311] "No that's it […] we're not working in the same way that the doctors are working, you know. In our minds we're not sort of sitting there timing it" [2316] [agreement and laughter] "Oh I am thoroughly unstiffened, no […] that's not the real world" [2316]*

These questions were not felt to capture the whole experience of stiffness.

> *"But we've also been talking about people who are stiff all day as well. So that would be a different question […] 'Cause my experience was you know, the sort of the fixed hands. And then it would go during the rest of the day. But other people are talking about that they've got it during the day. So that would be another question that you'd need to ask as well […] If you're asking people if they are stiff during the day and then asking if they are stiff when they wake up, is there extra stiffness when you wake up? They are two different questions" [2103]*

> *"I kind of find the whole 'how long does your stiffness last?' quite a difficult one […] if they're just looking at morning stiffness, then that doesn't capture the general on-going seizing up through the day stiffness, sometimes it does but quite often, well it doesn't at all and morning stiffness for me is mostly where my RA is in a flare or it's not well managed […] at the moment it's sort of fairly okay-ly managed so I'm not getting a lot of morning stiffness but I do seize up through the day" [2316]*

It was also highlighted that the duration of other symptoms was not questioned, therefore why was it relevant for stiffness.

> *"Because we don't ask 'how long does you pain last?' Do we? […] So why do we always need to say 'how long does your stiffness last?' You know if you take painkillers it doesn't always make it go away. It might ease it so moving certainly*

on the whole eases off the stiffness but we don't expect to say 'oh my pain went within ten minutes' or 'my pain went away within an hour" [2316]

Specifically in relation to MS it was again highlighted that stiffness was relevant at times of day other than just the morning.

"I think, now we're talking about it, we only talk about the morning because you ask us" [2101] [laughter] "That's right, that's right" [2105] "But it is, it's after anytime" [2101] "Any time of day really" [2104]

It was also suggested that questions about the morning make an assumption that people only get up in the morning and that they get up when they wake up. Therefore there was uncertainty about what information to include in answers to these questions.

"I mean, dependent on age and other things I mean we don't have an unbroken eight hour period [of sleep], which I think is what […] these young doctors think we do [laughs]" [2101]

Instead of asking about the duration of MS, patients suggested that asking 'when are you stiff' may be more relevant and would be easier to answer. Suggestions for formats for this were described and drawn up on the flip chart.

"[…] is there any aspect that particularly springs out when you think about assessing stiffness?" [Halls] "When is it worse throughout the day? And is it on waking, is it mid-morning, is it lunchtime, is it afternoon, is it when you feel you're tired? […] I think it's important that you know which parts of the day that individuals have the worst problems? [2208]

"I think you've got to differentiate, haven't you, between whether it's all day, morning, evening or whenever, really" [2209] [agreement] "You could do AM, PM, noon, sorry AM, PM, night" [2208] "Well, couldn't you have hourly […] how often does your stiffness affect you on a 24 hour basis" [2210] […] "I was wondering, could you have a clock" [2209] "with the hours and then some people might be able to shade it in [2208] "Yes, you could have an AM/PM clock" [2210]

It was also suggested that asking 'when are you stiff' would capture stiffness during the night which was relevant for some participants.

5.5.2.3.1.5 Location

The idea of the number of affected joints was originally identified in Study 1. It was proposed to groups for discussion as a topic that required further elaboration.

"So could we ask a question here maybe about if stiffness affects more joints or is more widespread, would that indicate it being worse?" [Halls] "It would definitely be worse" [2206] [agreement]

"[…] but when you're talking about levels, there could be just a question as to whether it's just one joint or you know, a number of joints. Say, is it one to three joints?" [2314] "OK so number of joints" [Halls] "Is it all over stiffness or is it […]?" [2316]

It was suggested that location was a natural way to think about the individual experience of stiffness. It was highlighted that this could be captured visually and potential formats were described and drawn up on the flip chart.

"You know the picture they have of a person […] with the massive hands […] I always go to my consultant, that's how it feels [laughs] […] it would be quite nice if you know, you could say, these bits are stiff" [2103]

It was also suggested that stiffness may affect different joints in different ways.

"I've just put up a sort of example, of a question […] What do you think of that one?" [Halls] "Again, it depends on the joint, doesn't it?" [2105]

"[…] can we ask what your level of stiffness is? [Halls] "But wouldn't that be like on a level of each different joint then no? […] Because I find the different joint like I mean obviously my wrists are stiff, they don't move so like up to my knees, which will bend and like my elbow. They don't bend straight out so they are like totally different so then obviously there's a level between each joint then isn't there" [2315] "That's another point yeah" [2314] "That's really interesting okay so in terms of severity we'd have to consider different joints? [Halls] "Yeah" [2315] "Yeah" [2313]

It was identified that asking about location also related the usual experience of stiffness for each individual.

"You'd want to say […] it's my fingers or my knees" [2103] "Yes, you've got to define the part of the body" [2105] "[whether] It's a new thing. You know?" [2103]

"[…] so have a way of saying okay my stiffness level that day is on this line but you know we're both saying I can mark or we can mark either the whole body as stiff or these joints are particularly stiff that day. Like you do with the inflammation. They tick off which joints are stiff yeah" [2316]

5.5.2.3.1.6 Pain, stiffness and other symptoms

As in Study 1, the relationship between stiffness and other symptoms was discussed. In relation to measurement, it was suggested earlier (Section 5.5.2.2.1) that questions about different symptoms needed to be asked separately for clarity. However, another

group suggested that the combination of stiffness and pain had more impact than just stiffness, thus it may be relevant to ask about both together.

> *"[…] do people need to ask about painful stiffness?" [ED] [agreement] "Yes. You've got to link the two together" [2210] "Because one goes with the other. So if you're stiff and you want to do something then you know you're going to get pain" [2206] "[…] so the question is, if that joint is stiff […] does it stop you doing something because it's painful, or because it's stiff" [2210] […] "As we said earlier, the stiffness is a warning that if you carry on […] It's going to hurt" [2206] […] "I don't think the stiffness has that major impact, but put pain with it, it becomes a major impact" [2210]*

Furthermore, stiffness as a result of different processes also complicated this issue as it was suggested that stiffness as a result of mechanical processes was not related to pain and therefore has less impact.

> *"[…] both my wrists are fused, so they're stiff. No pain, but they're stiff all the time, but I still get on with things and I haven't got pain, I can't bend them, right, I can't bend them or nothing like that, but they're stiff […] they're stiff, but I do everything, and I don't even give them a thought that they're stiff, because there's no pain with it. So take away the pain from stiffness […] it wouldn't affect me, would it" [2210]*

A method of capturing pain, swelling and stiffness was suggested by two groups relating back to the visual format proposed earlier (Section 5.5.2.3.1.5).

> *"If you are interested in measuring stiffness, well isn't the next question, is this accompanied by any pain? […] Quite happy to separate them out" [2101] "You could have two little men or women […] One for stiffness and one for pain" [2105] […] "you can say well […] I've got the stiffness but no pain" [2101]*

> *"[…] there's three questions there, isn't there, that you could ask and have tick boxes, and then underneath, how does each one affect […] your daily life? […] The stiffness, yes or no, with pain, yes or no, with swelling, yes or no" [2210]*

5.5.2.3.2 Response options

Response options were discussed mainly in focus groups 1 and 2. There were differences in opinion about response options. Some patients highlighted that VAS were easy to respond to but others felt they were imprecise.

> *"I don't like lines. Because they are so imprecise. And in my brain, boxes are more precise" [2101]*

> *"And those variation lines are good" [2210] "Yes" [2208] "What do you mean?" [Halls] "Like on there [VAS]" [2208] […] "Because all you've got to do is put a little line down, it's so easy" [2210] "That's right" [2206]*

However, other patients preferred NRS to VAS as the numbers made it clearer.

> *"I would prefer numbers, personally" [2208] "Yes, it is easier, 1 to 10" [2206] "1-10 rather than putting a line in […] I think a number is much easier, rather than a little line" [2208]*

Other patients suggested that NRS provide too many options.

> *"I have a problem over what's the difference between five and seven?" [2105] [agreement] "No tell me. What is the difference? Subtle difference, between five and seven. It's over complicated and it tells you nothing" [2105] "Oh, so what would you rather have?" [2103] "I would rather simply have low, medium and high. A, B, C. Whatever" [2105]*

When asked about a Likert scale response, some patients suggested a shorter option.

> *"If you said agree on that, they may ask you then, why do agree, and that goes on and on and on then, doesn't it?" [2206] "Yes or no then" [2210] "Yes or no would be better, yes" [2206]*

Fewer options were felt to reduce burden on the participant.

> *"If you had to do the degree [of stiffness], how would you like that?" [Halls] "1-10" [2105] "Yeah" [2103] […] "Or if you want to make it simpler, A, B, C […] None or moderate or severe. Three options, you know?" [2105] "Yeah yeah. Low, moderate and high or something" [2101] "Low, moderate and severe" [2105]*
>
> *"Yes, less options" [2206] "Yes, less options" [2210] "More straightforward questions, less options" [2206]*

5.5.2.3.3 Timeframe

There was considerable debate about the timeframe over which it would be relevant to ask about stiffness. It was felt that questions must ask about a recent timeframe.

> *"It's no good going over last year" [2208] "No" [2206/2210] "It's, it's it's recent, has to be pretty recent" [2208]*

Initial discussion suggested that stiffness over the last week would be acceptable and accurate.

> *"I think most people could remember how it's been over the last week" [2208] […] "I think you'd be able to remember the last week" [2209] […] "Because if you go*

too far it's just guess work" [2210] "That's right. And then it's not true […] It's not credible, is it" [2208]

However, it was also suggested that over the last week would not capture the variability of stiffness within recent months or within that week.

"You can remember and recall [a week] properly, accurately" [2208] "Do you think people could say over the last week, or do you think there'd be so much possible up and down over that week that…? [ED] "It varies every day, really" [2206] "That's true, isn't it, yes" [2209]

Another suggestion was current stiffness which was considered to be relevant and would provide a record if it was documented.

"And that's what you are now. Today. When they ask you. Over the phone or with a consultant" [2105] [agreement] "I think 2105 has used that expression, snap shot" [2101] [agreement]

However, it was also acknowledged that this approach may not capture stiffness variability either.

"If you think when you're stiff, the fact that you've got here quite often means you've walked off that stiffness or you've moved off that stiffness because you've got up, you've got yourself together, you've got on the bus or you've driven down. A lot of the stiffness that you'd have had in the morning, has gone because you're already here and you're moving so it's kind of, we need to have something that kind of captures what was it like overnight or that day" [2316]

5.5.2.3.4 Layout and format

The first aspect that was discussed in relation to layout and format was that images may be an appropriate way of capturing stiffness information. One suggestion was using a visual clock face to record when participants are stiff.

"I was wondering, could you have a clock" [2209] "Yeah yeah" [2210] "with the hours and then sort of, some people might be able to shade it in" [2208] "Yes, you could have an AM/PM clock" [2210]

Another suggestion, discussed in two focus groups was using a diagram of a person to indicate the location of stiffness. It was felt that this approach would be particularly effective at capturing the individuality of the experience of stiffness.

> *"A person, yes, a little pin figure" [2209] "The only trouble with that is you'd be ticking it all over" [2206] "But not all of us, we wouldn't all be ticking it all over" [2209]*

Another point was the importance of simplicity of the questionnaire.

> *"Which is why it needs to be fairly simple doesn't it" [2316] [agreement] "It can't be too complicated" [2316]*

This specifically related to preferences for fewer questions and response options to enhance simplicity.

> *"Yes, less options" [2210] "More straightforward questions, less options" [2206]*

It also related to the use of current language to aid clarity and understandability.

> *"And try and update these very old fashioned questions" [2101] [laughter] "It really does makes me feel, I am stereotyping, a little old lady. Well, I am a little old lady. But, you know, doddering around you know with the bath and no shower" [2101] [laughter]*

There was also discussion regarding the format of the response options. One group suggested that free text options would allow appreciation of the individuality of the experience of stiffness.

> *"[…] maybe room for the person to write in their own remarks as well, depending on what they've [put], can you explain this […]" [2207]*

However, it was also recognised that free text may not be appropriate for generalising responses.

> *"[…] you need to generalise about how it affects you […] because otherwise, you start listing things you could go on forever about [how] it affects you" [2210]*

Practical considerations regarding marking responses were also highlighted. This included that a simple mark is preferable to considerable writing and that circles are more difficult than crossing a box.

> *"I think things to circle or tick is better because if you've got to write it, sometimes it's very difficult" [2206] "Sometimes you can't write" [2208]*
>
> *"[…] I don't like circling […] cross out is best" [2210]*

A final interesting point was made in two focus groups that indicated that participants would like to know the purpose of the questionnaire and what the answers you give is going to affect.

> *"Yes, but as you said if we know why you want to measure in a particular way […] it's fine […] But it's not knowing how that's going to affect how you are going to be treated. Whether it's for you, or for the practitioner" [2101]*

This related to a point made by one participant in focus group 3 who questioned the purpose of completing a pre-clinic questionnaire, again stressing the importance of ensuring patient awareness of purpose to ensure accurate results.

> *"But then I don't know, does anybody actually look at those things?" [2315] [laughter] "She rubs them out doesn't she!" [2313] "Yeah but that's what they do. I say that because I fill them out all the time, yeah and I just go, yeah, yeah [demonstrates ticking each box not looking] […]" [2315]*

Finally there was brief discussion about how to word items. This included whether to include 'stiffness' in every stem question (focus group 2) and how to ask about severity (focus group 3).

> *"[…] would it be helpful to have the word stiffness in the questions […]" [ED] "Yes, but not in every one […] although, I suppose because it's about stiffness then it needs to be kind of […] affirmed" [2207]*

> *"I don't know, yeah I would say level, I would say level yeah" [2315], "I would say level" [2313] […] "So stiffness level, like pain level" [2316] "So if we asked you a question about your level of stiffness?" [SH] "Yeah" [2313]*

5.6 Discussion

Focus group data supported the Study 1 results, and appeared consistent with the experiences of this new population of patients, using a different method of data collection. This enhances the robustness of the stiffness conceptual model. In addition, a significant new finding was the information gathered regarding the patient perspective of stiffness assessment. This provided insight into patient relevant stem question categories and preferences regarding response options, timeframe, and format.

The consistency between the results from Study 1 and Study 2 is reassuring and demonstrates the strength of the findings (Mays and Pope, 1995). Although additional codes were added to the coding framework during analysis of Study 2, these

elaborated areas that had been identified in Study 1 rather than generating new and unexpected areas of discussion, confirmed by no new themes being identified. Therefore the conceptual model appears to provide a robust foundation of qualitative data on which to base and develop items for a stiffness PROM.

In relation to the patient perspective of stiffness assessment, a number of key areas were identified. The first related to the identification of difficulties with, and dislike of, the concept of duration. Given the variations in the topic guide between focus groups, this was discussed in more detail in focus group 1, but interestingly was also identified in focus groups 2 and 3. Particular issues with duration questions included difficulties regarding quantification, uncertainties regarding the start or endpoints of such questions, and concerns that the concept does not capture the full stiffness experience. Difficulties with the wording of duration questions in other literature have been highlighted previously (Section 2.4.1.4) and is concerning given that MS duration is the most commonly employed stiffness assessment method in trials (Kalyoncu *et al,* 2009). However, Study 2 did provide information that may allow development of the concept of duration to be more relevant to patients with the suggestion that assessing the temporality of stiffness was still important and may capture more of the patient experience of stiffness. Study 2 data also reinforced the relevance of the concept of impact. The participant emphasis on impact was a finding that emerged strongly from Study 1 and has been identified in other research as a topic of shared relevance across participants (Orbai *et al,* 2014). The relevance of the impact triad to patients is unsurprising as it was developed in collaboration with patients to capture patient relevant outcomes (Sanderson *et al,* 2011). However, as it is recommended for consideration in the development of PROM (Sanderson *et al,* 2011), it was relevant to explore in this study and take forward to later stages of item development. Another key finding related to the wording used to describe the experience of stiffness. Within the literature, many different words are used to describe the experience or sensation of stiffness (e.g. Helliwell, 1995) and when proposed to participants in this study, many of these words were relevant, although their use was not often initiated by participants. The term *stiffness* had been used consistently throughout Study 1 and Study 2 by participants and the focus groups provided clarification that it was a patient relevant word rather than a medical term used by clinicians. The many words discussed in this study and other literature may be useful in certain contexts, for example in clinical consultations where the nature of the stiffness experience is being discussed. However, in relation to the development of a stiffness PROM it appears that the word *stiffness* would be an appropriate and

acceptable term to use. The term *stiffness* is also consistent with the patient descriptions from Study 1 and 2 that the symptom is not only experienced in the morning period. This challenges the traditionally accepted concept of MS or EMS and was also supported in the work by Orbai *et al.* (2014).

Consistent with discussions in Study 1, the close relationship between stiffness and pain was discussed across all focus groups, particularly in relation to the specificity of stiffness. Interestingly, the researcher felt that this relationship was much more strongly conveyed by participants during Study 2 compared to Study 1. This may be a result of differences in the methods of data collection. It is suggested that interaction between participants generated in focus groups can lead to reinforcement of similarities and differences (Lambert and Loiselle, 2007). As pain has been suggested to be a priority for people with RA (Minnock *et al,* 2003), it is not surprising that it was relevant within discussions. Given the opportunity to identify similarities in shared experiences from the focus group method, and the prominence of pain in the experience of RA, this relationship may have been accentuated during Study 2. As discussed earlier, other research has suggested stiffness-pain interdependence based on the results of a qualitative focus group study (Orbai *et al,* 2014). In contrast, although Studies 1 and 2 provide evidence of a close relationship between these symptoms, they also indicated that stiffness can be separated from pain by impact, and differences in management strategies and medications. Furthermore, some participants identified difficulties when responding to questions asking about pain and stiffness together in one question. Overall, differences in the emphasis placed on the relationship between pain and stiffness, in otherwise very similar pieces of work (Orbai *et al,* 2014; Halls *et al,* 2015) may in part be explained by the data collection method. Yet, both studies conclude that stiffness is an important aspect in the patient experience of RA and that the development of an appropriate assessment tool for stiffness specifically is significant (Orbai *et al,* 2014; Halls *et al,* 2015). The identified relationship between pain and stiffness does have implications for item development where it will be important to reinforce the emphasis on stiffness specifically (e.g. state in stem questions), and also base measurement on concepts that enable differentiation (e.g. impact).

A final discussion point relates to the participant identification of stiffness in diseases external to RA. This is not surprising as stiffness is a common symptom in other rheumatic conditions. However, what was interesting was the suggestion that there may be similarities between those experiences. While some participants in the focus

groups felt that the experience of stiffness was different across conditions, others identified that it was similar. This was consistent with findings from Study 1 and is particularly relevant in relation to stiffness assessment in both an RA context and more broadly in the wider rheumatology community. If there are tangible similarities in stiffness experiences across conditions, this may have implications for the wider use of any new measurement instrument developed in an RA context. The value in exploring stiffness assessment across conditions has been recognised with the development of an OMERACT special interest group to coordinate further research in this area (Orbai *et al,* 2015).

The total sample in Study 2 was small (n=16) but each focus group included participants with a range of age, gender and disease duration, and was similar to the Study 1 sample. The main difference between the samples was that all but one participant was recruited from a single site (BRI) as recruitment at NBT was stopped due to other research commitments and considerations regarding the provision of focus groups in convenient locations for all participants. Therefore, the sample may represent a slightly less diverse population. Another consideration related to the purposive sampling approach, which was employed to enable recruitment of a range of participants. However, in practice it proved difficult to systematically reflect a range of characteristics within each focus group session while ensuring recruitment of enough participants per group. This practical consideration meant that although each group included a range of participants, this is only partly due to the sampling approach, which was more of a convenience sample than the intended purposive sample. Another limitation relating to recruitment was the poor response rate (15%). It was likely that recruitment was affected by being performed during the Christmas and New Year period. It also may have been influenced by different researchers recruiting for this study and other studies simultaneously which may have diluted the emphasis on all studies. Despite these considerations, the use of triangulation to validate qualitative results (Ritchie, 2003) demonstrated the consistency of results across Study 1 and Study 2, which is a key strength of this study. The reliability of findings was enhanced by discussion during analysis with other members of the supervisory team (Mays and Pope, 1995; Cohen and Crabtree, 2008). Furthermore, Study 2 was reported in accordance with the COREQ framework to enhance transparency of the qualitative work (Tong, Sainsbury and Craig, 2007) (Appendix L). The final consideration in this study was the influence of the progression exam, which resulted in changes to the focus group topic guide between focus group 1 and focus groups 2 and 3, which fits with recommendations that topic guides are refined and

revised throughout the process of qualitative data collection (Arthur and Nazroo, 2003). Furthermore, as the aims and objectives of the study remained consistent and the approach to analysis was not affected (as it may have been in an analysis such as content analysis), it could be argued that changes made in order to collect data more effectively is a strength of this study. Overall, although the feedback from the progression exam had time and content implications, it enabled the researcher to reconsider aspects of the study and be in a better position to capture relevant information.

5.7 Conclusion

This study has supported and reinforced the conceptual model identified in Study 1 in a new population of patients, using a different method of data collection. It has also provided information that specifically addresses stiffness assessment from the patient perspective. A number of concepts for measurement instrument development were proposed for further exploration including individuality, impact, temporality, immobility, and location. Patient preferences to capture and format these concepts were discussed. These patient-driven concepts now require consideration alongside measurement theory to develop a conceptually sound yet practically appropriate draft set of items. Chapter 6 will describe the process of combining the qualitative data generated in both Study 1 and Study 2, along with measurement theory evidence, to inform item development.

Chapter 6: Item development

Preceding chapters have developed understanding of the patient experience of RA stiffness and stiffness assessment from the patient perspective. This chapter will now describe the development of preliminary items for an RA stiffness PROM using these data. The chapter will discuss how relevant aspects for measurement were identified from the codes and themes in the conceptual model, and were combined with the stiffness assessment data. It will illustrate the process of drafting and re-drafting of items, and the involvement of the supervisory team.

6.1 Background

As highlighted previously, current stiffness assessment relies on non-standardised EMS/MS duration or severity questions which are poorly defined, have limited measurement property evidence, and do not appear to have been developed according to current standards including collaboration with patients (USDHHS FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b) (Chapter 2). Difficulties when responding to currently used duration questions have been identified (Section 2.4.1.4 and Study 2), and consideration of more appropriate and patient relevant concepts, such as impact, has been suggested to assess stiffness (Studies 1 and 2; Orbai *et al,* 2014; Halls *et al,* 2015; Orbai *et al,* 2015). Given this, the development of items for a new RA stiffness PROM that capture the patient perspective and demonstrate face and content validity is necessary.

Face and content validity refer to whether a measure looks appropriate (Streiner and Norman, 2008), and are both essential in PROM development (Frost *et al,* 2007). Studies 1 and 2 provided qualitative data to inform the development of items for new RA stiffness PROM. Therefore, item development was based on the stiffness conceptual model, and coding trees generated in earlier qualitative studies (Chapters 4 and 5). It was important that the patient-driven concepts generated in the qualitative data were considered alongside measurement theory to develop conceptually sound yet practically appropriate preliminary items for use in clinical and research environments. Consideration of PROM development theory (Section 2.3), such as the OMERACT Filter which was developed to determine the applicability of measurement instruments (Boers *et al,* 1998), was considered throughout.

6.1.1 PROM development guidelines

The USDHHS FDA guidelines (2009) concentrate on the development of PROMs specifically for the purpose of supporting pharmaceutical labelling claims. The ISPOR guidelines (Patrick *et al,* 2011a; Patrick *et al,* 2011b) focus specifically on the development of content validity in PROM development. Although the primary reason for the development of this stiffness outcome measure is not to support labelling claims, it is an area that pharmaceutical companies might wish to explore in future work and as such guidelines provide a rigorous and systematic process to guide PROM development, they were used to inform the development of preliminary items. The generation of items should include involvement from the relevant population, as should the development of appropriate item wording and also assessment (USDHHS FDA, 2009). This chapter addresses preliminary item wording, stem question and anchors, response options, timeframe, and format by developing them from the stiffness conceptual model and coding trees generated in Chapters 4 and 5, in consideration with measurement theory.

6.2 Aims and objectives

The aim of this study was to use the qualitative data to develop preliminary content for an RA stiffness PROM. The study objectives were:

- Demonstrate the process of combining the qualitative data from Study 1 and 2 and the drafting and re-drafting of items
- Develop a list of preliminary items using the concepts and language identified by patients in Study 1 and 2 including appropriate instructions, wording, response options or anchors, timescale, and format
- Prepare the agreed preliminary items for cognitive interviewing with RA patients (Study 3)

6.3 Item development

Item development was performed between February and June 2014. It was implemented from the dual perspectives of being informed by the qualitative data generated in Studies 1 and 2, measurement theory (Section 2.3), and consideration of the purpose that the developed tool would serve. Item development involved an iterative process of development and discussion with different members of the supervisory team.

It is suggested that the source of items and their subsequent editing, selection, and reduction should be documented (Patrick *et al,* 2011b), which can be performed using tracking tables (Patrick *et al,* 2007; USDHHS FDA, 2009). The development of items is described below and illustrated in a series of tables (Appendices E, J, K, M, N, O, P) starting from the coding tree identified in Study 1 (Appendix E), and ending with the final set of items to be taken to cognitive interviews in Study 3 (Appendix P).

6.3.1 Moving from the qualitative experience of stiffness to stiffness measurement

Studies 1 and 2 provided insight into the patient experience of stiffness and stiffness assessment. This broad information, particularly from Study 1 was vital in gaining understanding into the patient experience of a seemingly complex symptom. However, some of this information was more relevant for experiential understanding than for measurement purposes. Specifically for this project, the aim was the development of a stiffness PROM for use in clinical or research situations. Use of PROMs in these areas include testing the outcome of a treatment or intervention (research), or monitoring patient progress (clinical) (Nelson *et al,* 2015), as captured in the OMERACT Filter heading of discrimination (Boers *et al,* 1998). Thus, it was important to identify aspects that address these purposes. Additionally, although the data gathered regarding the experience and assessment of stiffness were analysed separately in Study 2, participants discussed the two issues in the context of one another. Therefore, to retain that context, information from both areas of analysis were mapped on to each other to proceed with item development.

Firstly, the coding tree of patient experiences of stiffness generated from Study 1 (Appendix E), expanded with data generated in Study 2 (Appendix J) was reviewed and discussed with the supervisory team to identify aspects only relevant to measurement (Appendix M). Appendix M demonstrates that three groups of codes were removed as they were only relevant to the development of experiential understanding of stiffness: varying prominence of stiffness across the course of the disease (level 2 group); medications have other considerations (level 1 group); and internal – part of general RA management (level 1 group). All other information was retained and renamed to clarify the target for measurement. Figure 6.1 provides a concise account of the key areas and associated level 2 groups (renamed for clarity) that were identified for potential measurement.

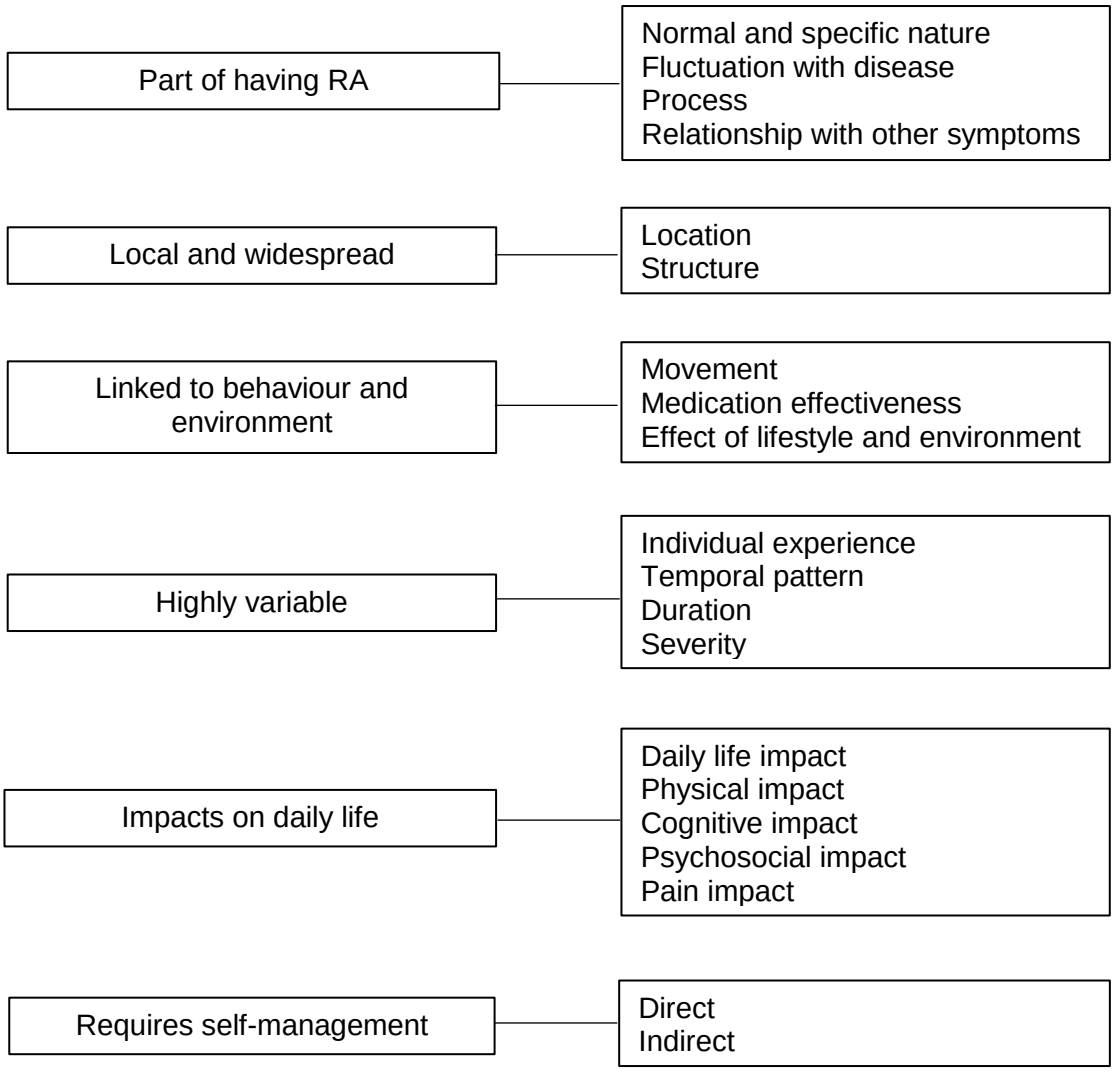


Figure 6.1: Key areas for measurement from the conceptual framework

Secondly, review of the data relevant to stiffness assessment generated in Study 2 was performed following the same process. Again, this was examined to identify key aspects for measurement which were grouped together and renamed (Appendix K, column 5).

Finally the data regarding the experience of stiffness (Appendix M) and stiffness assessment (Appendix K) were mapped onto each other to combine all key data from all sources (Appendix N). The first column in the Table in Appendix N lists the key areas identified from the experience of stiffness (Appendix M, column 6). The second column includes the key areas regarding stiffness assessment (Appendix K, column 5) which were mapped onto relevant areas from column 1. Columns 3 and 4 provide a summary of the ideas for measurement and points from discussion with supervisory team members.

6.3.2 Process of item development

The retained and reorganised data were then used in the development of a set of preliminary items. Initially this focused on development of stem questions and anchors with all potential item ideas within each key area for measurement being documented and developed though discussion with the supervisory team. Appendix O details the earliest phase of item development and includes the key areas for measurement (column 1), the potential items to capture within these (column 2), draft wording (column 3), and considerations from discussion with the supervisory team (columns 4 and 5).

Following this, a first draft of all preliminary items was developed. Appendix P is an item tracking table including each preliminary item, revision and development points from discussion with supervisory team members, and the final items to be taken forward for testing in cognitive interviewing (Study 3). This process involved consideration of each of the components of item development; stem questions and anchors; response options; timeframe; and layout and format.

6.3.2.1 Stem questions and anchors

The stem questions and anchors were formed from the qualitative data generated in Studies 1 and 2 together with consideration of ideas or concepts within the pre-existing literature. Preliminary item stem question and anchor development is described under five sections; 1) severity items; 2) impact items; 3) attribution items; 4) traditional stiffness items; 5) response shift items but were not defined a priori. Each section is described individually below along with any relevant literature. The sections also map onto the item tracking matrix (Appendix P) where the 'item section' collumn highlights the relevant section for each preliminary item.

6.3.2.1.1 Severity items (Items 1-7, Appendix P)

Patient relevant concepts to address severity were identified from earlier qualitative studies (Chapters 4 and 5). Each has been described below along with any relevant literature.

6.3.2.1.1.1 Location

During qualitative studies (Chapters 4 and 5), location was identified as a way of discussing stiffness that was relevant to the individual experience. Patients in Study 2 felt that location was an intuitive way of being asked about stiffness and suggested

it could be captured using a diagram. This was explored during early item development (Appendix O).

Within the literature there are other assessments that employ the use of diagrams, including those used in the assessment of pain. The Brief Pain Inventory (developed for use in patients with cancer) asks responders to mark the location of their pain on front/back body diagrams (Cleeland and Ryan, 1994). The Body Chart (developed for use in AS) goes one step further by asking responders to mark the location of their pain and its severity on a 4-point scale (1=mild, 4=very severe) (Dziedzic, 1997). Within the RA literature, the use of diagrams in assessment are employed as part of questionnaires assessing patient reported disease activity, including the RADAI (Stucki *et al,* 1995) and the PDAS2 (Choy *et al,* 2008). These assessments include diagrams for patient reported tender (RADAI, Stucki *et al,* 1995) or swollen (PDAS2, Choy *et al,* 2008) joint counts but there are no examples of stiffness assessment using diagrams. When the idea of the use of a diagram was discussed with one patient partner (GB) it was suggested that this would be possible for patients to complete, but that it might be optimal to provide a diagram and a written question so that patients could complete whichever version they found easiest. Additionally, when attempting to design an item using a diagram, a number of practical issues were identified, including how to instruct responders to mark it, and how it would be scored. Although there are options regarding the scoring of diagrams including the use of percentage estimates (e.g. Margolis, Tait and Krause, 1986; Margolis, Chibnall and Tait, 1988), and more recently using computer software (e.g. Jaatun *et al,* 2015), feasibility in clinical or research settings was questioned. Concerns relating to the OMERACT Filter component of feasibility (Boers *et al,* 1998), and the suggestions from one of the teams patient partners (GB) led to the decision to develop this as a written question rather than as a diagram.

6.3.2.1.1.2 Timing

The temporal pattern of stiffness was an important area for patients. Study 1 identified that the experience of stiffness was not limited to the early morning, on which much of traditional assessment is based and although in Study 2 patients discussed that the traditional duration items were hard to answer, the timing of stiffness was still felt to be important. To assess timing, a question was generated to capture when stiffness occurs. This removed the traditional emphasis on EMS/MS that patients disliked and felt didn't fully capture their experience. However, that the stiffness might have occurred in the morning was available as a response option, thus capturing

information relating to current understanding of stiffness pathophysiology. As above, this item could have been captured in diagram format (e.g. clock face) but for practicality reasons it was developed as a written item.

An item to capture increased stiffness variability was also generated. This aimed to capture the patient suggestion from Study 1 that during higher disease activity, stiffness was more variable (e.g. more frequent in occurrence).

6.3.2.1.1.3 Stiffness after immobility

Stiffness after a period of immobility was another idea generated from Studies 1 and 2. Interestingly the definition resulting from previous qualitative work on stiffness in RA included stiffness 'after immobility' but confusingly this was included in the definition of MS (Lineker *et al,* 1999). Furthermore, two of the 36 individual stiffness assessment measures identified in the systematic literature review update investigated this concept. One article assessed 'starting stiffness after a time of rest' but did not define the wording used (Leeb *et al,* 2003), likely because it was developed in German and had not been validated in English (Rintelen *et al,* 2009). The other assessed stiffness using the question 'How severe has your stiffness been after sitting or lying down or while resting later in the day?' (Wolfe, 1999). As this item had been developed as part of the WOMAC in an OA population (Bellamy *et al,* 1988) it was felt that the development of a new item using data collected from an RA population would be appropriate.

6.3.2.1.1.4 Medication effectiveness

Discussion regarding medications was identified as relevant to patients during Studies 1 and 2. A preliminary item was based on an item used in the AS literature that is sometimes included above the BASDAI (Garrett *et al,* 1994) which asks responders to indicate the effectiveness of their medication in relieving symptoms on a VAS (0=No effect, 10=Very effective).

6.3.2.1.2 Impact items (Items 8-59, Appendix P)

The importance of the impact of stiffness to RA patients was highlighted throughout Studies 1 and 2. For a symptom with such variability and individual nature, impact has been described as a "common language" across RA patients (Orbai *et al,* 2014, p.10). For this reason, items to capture the impact of stiffness were developed in two ways: 1) impact items developed directly from the qualitative data generated in earlier

studies (Section 6.3.2.1.2.1); 2) impact items developed based on the concept of the impact triad (Section 6.3.2.1.2.2). These are described below.

6.3.2.1.2.1 Impact items generated directly from qualitative data

A number of items to capture impact were developed based on the combined coding framework from Studies 1 and 2 (Appendix J) and the patient language used in those studies. It has been suggested that outcome measure development can be informed by the conceptual framework of the international classification of functioning, disability and health (ICF) and its disease specific core sets (World Health Organisation, 2002). The ICF is a framework providing a standard approach to describing health and health related conditions (World Health Organisation, 2002). ICF core sets have also been developed for specific conditions including RA. The ICF core set for RA is an RA specific framework that includes 96 ICF categories in four sections: 1) body functions (e.g. mobility of joint functions); 2) body structures (e.g. elbow joint); 3) activities and participation (e.g. doing housework); 4) environmental factors (e.g. immediate family) (Stucki *et al,* 2004). However, the ICF core set for RA was not used as framework to inform the development of the stiffness items because it captures all aspects of the experience of RA for patients (Stucki *et al,* 2004) making it very broad and likely including aspects that are not relevant to stiffness specifically. Furthermore, although integration would have been possible, not using the ICF core set for RA framework meant that item development could be based on the patient-generated stiffness conceptual model identified in Studies 1 and 2, which reinforced the content validity of item development.

6.3.2.1.2.2 The impact triad

One concept that related specifically to aspects already captured in the literature was the impact triad. The impact triad was developed by patients and researchers and recommends considering not only the severity of an outcome, but also its importance to patients, and their ability to self-manage it (Sanderson *et al,* 2011). These three aspects combine to form impact. The concept of the impact triad has been recommended for inclusion in the development of PROMs, particularly given its relevance from the patient perspective (Sanderson *et al,* 2011). As the best approach to its assessment is currently undetermined (Sanderson *et al,* 2011), the approach used in the development of impact triad items was based on a previously used method, employed in the development of the Bristol Rheumatoid Arthritis Fatigue NRS (BRAF-NRS, Nicklin *et al,* 2010a; Nicklin *et al,* 2010b). This involved developing

one question for each component of the impact triad. In discussion with the supervisory team, some of whom had been involved in the development of the impact triad, it was also recommended that a question be developed to capture the overall concept of impact.

In relation to the development of items to capture the impact triad, the RA literature includes a considerable amount of research into self-management or self-care. Self-care is defined by the Department of Health as "[…] the actions people take […] to stay fit and maintain good physical and mental health; meet social and psychological needs; prevent illness or accidents; care for minor ailments and long-term conditions; and maintain health and wellbeing after an acute illness or discharge from hospital" (Department of Health, 2005, p.1). Some suggest a distinction between self-management and self-care with self-care being a 'normal activity' relating to daily lifestyle decisions, and self-management being an extension of that, relating specifically to "ailments" (Chambers, 2006, p.129; Ahmad *et al,* 2014). As defined earlier (Section 4.6), the term self-management is used according to Barlow's (2001) definition which captures the broad descriptions of self-management used by patients in Studies 1 and 2. Specifically in an arthritis context, self-management has been identified as important for patients in recent qualitative work (Ryan *et al,* 2013), and also as a component of the impact triad (Sanderson *et al,* 2011). Different words to capture the patient perception of their ability to self-manage have been used in the literature. During the development of the BRAF-NRS, questions using both 'cope' and 'manage' were tested and patients suggested a distinction between them where 'manage' was more practical and 'cope' was more relevant to emotions. Although patients did also use the terms interchangeably. The final BRAF-NRS included the 'cope' wording (Nicklin *et al,* 2010b). In the qualitative data gathered in Studies 1 and 2, words including 'cope', 'manage', 'adapt', 'deal with' were used by patients. When discussed with the team, a patient partner (PR) identified that the term 'self-management' was not patient-driven and instead suggested 'cope', 'deal with', and 'make do' as more patient relevant. Given the lack of consistency in the wording to capture 'self-management', several items using different words were taken for testing in Study 3 to further explore the patient perspective regarding the most appropriate wording.

In addition to discussion regarding stem question wording, in the BRAF development there was also discussion regarding the anchors in the 'cope' question (Nicklin, 2009). Often in the development of scales, positive anchors (doing well, score 0) are placed

on the left and negative anchors (doing badly, score 10) on the right (Meyer, 2007). In the development of the BRAF-NRS, the patients in the focus groups felt it intuitively better to put the coping anchors the other way round with not coping (0) on the left, and coping well (10) on the right hand side of the scale. Although there were differences in patient opinion, overall patients felt that this was the best layout (Nicklin, 2009). Given the considerable patient involvement in the development of the BRAF (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b), the same anchor placement was used for this item in this study for the coping items.

6.3.2.1.3 Attribution items (Items 60-68, Appendix P)

Some items were developed in an attempt to capture the different ways in which patients described the experience of stiffness. These included perceptions that stiffness was influenced by other symptoms, disease activity, joint damage, the weather, doing too much or too little, and the effect of medications. There was discussion with the supervisory team about the relevance of these items from the perspective of their use in a measurement context as they seemed related to stiffness characterisation rather than quantification. But it was felt that if this was the way that some patients made sense of stiffness they should be tested as questions. These items were therefore included and taken for further investigation (Study 3).

6.3.2.1.4 Traditional stiffness items (Items 69-74, Appendix P)

It was also important to decide which traditional stiffness items from the literature to include in further testing. Traditional stiffness items assess a narrow range of concepts (MS or EMS, severity or duration) and are poorly defined or use variable wording or formats. There is limited evidence regarding the measurement properties of these items, and difficulties have been identified with some of the concepts assessed. The aim was to include a range of traditional stiffness items including those with defined wording, most commonly used, and with a range of different formats (Table 6.1).

Three traditional stiffness duration items were included (items D, E and F, Table 6.1). A recent review of stiffness assessment in low-disease states (van Tuyl, Lems and Boers, 2014) identified two stiffness PROM validation studies (Hazes, Hayton, and Silman, 1993; Khan *et al,* 2009) and those made conflicting recommendations regarding whether severity or duration was most effective. However, others have suggested that severity items are more effective than duration items (e.g. Westhoff *et*

al, 2008). Despite this uncertainty, duration items are the most commonly used stiffness assessment question in clinical trials (Cutolo, 2011), therefore it was important to include some duration items. Traditional stiffness duration item D is a component of a number of validated composite scales to assess disease activity, including the RADAI (Stucki *et al,* 1995; Fransen *et al,* 2000; modified from the RADAR (Mason *et al,* 1992)), and the PDAS2 (Choy *et al,* 2008; Choy *et al,* 2015). These scales have been validated in an RA population, and face validity was assessed as part of the development and validation of the PDAS2 (Choy *et al,* 2008). Traditional stiffness duration items E and F were both based on the same item wording. In a study that tested different wording of duration items, it was concluded that this wording was the best indicator of MS duration (Hazes *et al,* 1994). The wording of "morning stiffness […] to maximal improvement" is also in accordance with the ARA criteria (Arnett *et al,* 1988, p.315) and has been used in a number of other studies (e.g. Khan *et al,* 2009).

Table 6.1: Traditional stiffness items

Item concept	Item
Traditional stiffness severity item A	How would you describe the overall level of morning stiffness you have had from the time you wake up? 11-point NRS (0=No stiffness, 10=Very severe stiffness)
Traditional stiffness severity item B	How would you describe the overall level of morning stiffness you have had from the time you wake up? 100mm VAS (0=No stiffness, 10=Extreme stiffness)
Traditional stiffness severity item C	How would you describe the overall level of morning stiffness you have had from the time you wake up? 5 option adjectival scale (No stiffness, Mild stiffness, Moderate stiffness, Severe stiffness, Very severe stiffness)
Traditional stiffness duration item D	Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last? Less than 30 minutes, 30 minutes to an hour, 1-2 hours, 2-4 hours, More than 4 hours but less than all day, All day)
Traditional stiffness duration item E	How long does your morning stiffness last from waking until maximum improvement occurs? 3 option adjectival scale (Up to 1 hour, 1-3 hours, More than 3 hours)
Traditional stiffness duration item F	How long does your morning stiffness last from waking until maximum improvement occurs? Minutes/Hours

The timeframe over which stiffness is assessed is often not specifically stated in the wording of traditional items (e.g. Vliet Vlieland *et al,* 1997). Where a timeframe is specified, it is not standardised, for example, in the work by Hazes *et al.* (1994), responses to the stiffness item were collected in a daily diary, while in the study by Khan et al. (2009) the same wording was used but with a timeframe of a week. As stated in their paper, this may have been for consistency across all collected self-reported information (Khan *et al,* 2009). However, this lack of standardisation makes comparison across studies difficult. To test items in accordance with the literature where currently articles either do not define a timeframe or pick the most appropriate timeframe to suit their study, the wording of this traditional item in our draft PROM was included as explicitly stated, i.e. without a timeframe.

The response options provided for traditional stiffness duration items also vary, hence our duration items E and F differ. Item E provides responses on a 3 option adjectival scale. This was decided upon because fewer options and simplicity of items had been suggested by patients in Study 2, these categories had also been used in categorisation of responses in the literature (Hazes *et al,* 1994), and they provided fewer response options than provided in item D. Finally, item F reflected the commonly used response option provided for this item of minutes and hours (Rhind, Unsworth and Haslock, 1987; Arnett *et al,* 1988; Hazes, Hayton, and Silman, 1993; Hazes *et al,* 1994; Vliet Vlieland *et al,* 1997).

Three traditional unvalidated stiffness severity items were included (items A, B and C, Table 6.1). Despite stiffness severity items not being as common as duration items, some literature has suggested that stiffness severity items have better measurement properties than duration items (Section 2.4.1.4). Traditional stiffness severity items A and C were based on the items used in previous literature (Rhind, Unsworth and Haslock, 1987; Hazes, Hayton, and Silman, 1993). These studies explicitly defined the response options of the items used, but not the wording of the question. When looking across the literature to identify appropriate wording used for severity items, explicit wording could not be found within the RA literature. The only explicit wording identified was the wording of the severity item used in the BASDAI (Garrett *et al,* 1994), which although not ideal given its development and use in an AS population, provided defined wording, and used a VAS (item B). This question has also been used in studies involving RA patients (e.g. Lie *et al,* 2014).

6.3.2.1.5 Response shift items (Items 75-77, Appendix P)

Response shift has been defined as “a change in the meaning of one's self-evaluation of quality of life as a result of changes in internal standards, values and the conceptualisation of quality of life” (Sprangers and Schwartz, 1999, p.1509). The response shift theoretical model proposes how the response shift occurs and includes five key factors; a catalyst (e.g. change in health status), antecedents (e.g. gender), mechanisms (e.g. social support), response shift, and perceived quality of life (Sprangers and Schwartz, 1999; Schwartz and Sprangers, 1999). From a measurement perspective, it has been suggested that measurement may be influenced by the response shift and that assessment of response shift may be necessary to capture these changes (Sprangers, 2010). Interestingly, the patient data gathered in Studies 1 and 2 highlighted the individuality of the patient experience of stiffness yet how stiffness was a normal symptom of RA. This led to the suggestion to capture stiffness that was different (e.g. more severe) than RA stiffness that occurred most days and was now considered ‘normal’ by patients following a response shift in internal standards. Following discussion of this suggestion with the supervisory team, one item was developed in an attempt to assess response shift. This item would be used with the aim of standardising responses and would be placed within the demographics section so as not to influence response to other items.

6.3.2.2 Response options

It is recommended that the purpose and intended use of any PROM are considered in relation to decision making regarding response options (USDHHS FDA, 2009). As such, response options were a key component of draft item development. The current literature regarding response options relates to a number of areas including the optimal number and format of response options. It was also relevant to consider currently used response options in relevant literature and data capturing the patient perspective.

Firstly, the optimal number of response options was considered. The traditional work of Miller (1956, reprinted 1994) suggested that seven (plus or minus two) is the limit of the working memory thus that guideline has been implemented in a range of areas including response options in questionnaire development. Streiner and Norman (2008) suggest two further considerations related to this rule. Firstly, more response options may help to deal with ‘end-aversion bias’ where responders avoid scoring in the most extreme response categories. Secondly, when the aim is to gain a total score

from a number of individual items, reducing the response options to three or five will likely not result in loss of information. The latter is also consistent with earlier work that suggested four or five response options are the preference of survey scientists and reviewers because with fewer response options, there is less burden on the responder but still enough response options to be precise (Fries *et al,* 2006). These suggestions also fit with the patient preference for fewer responses and simplicity of items (Study 2). Following discussion with the supervisory team and in accordance with recommendations (USDHHS FDA, 2009; DeVellis, 2012), it was felt that primarily response options must fit the stem question. However, it was not vital or possible for all items to have the same uniform response options. This is also true of other PROMs including the BRAF (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b) and the HAQ (Fries *et al,* 1980).

Secondly, format of the response options was considered. There are many different available response option formats including Likert scales, VAS, NRS, and adjectival scales (Streiner and Norman, 2008; DeVellis, 2012). Each of these options has been described below along with relevant literature regarding the strengths and limitations of each approach.

6.3.2.2.1 VAS

VAS are commonly employed in medical contexts, especially in the assessment of pain (Huskinsson, 1974), but also in the RA specific context of disease activity assessment using the DAS28, which includes a PtG VAS (van der Heijde *et al,* 1992). VAS benefit from being simple, requiring only a mark on a 100mm line (Streiner and Norman, 2008). However, it has been suggested that while VAS may appear simple to researchers, this is not always the case for responders. In qualitative interviews with patients with a range of chronic conditions, VAS were the least preferred response option format compared to NRS and adjectival scales (Quadri *et al,* 2012). Similar findings resulted during the development of the RADAI questionnaire, where the VAS format used in the original RADAR questionnaire, from which the RADAI was developed (Mason *et al,* 1992), was changed to NRS because patients had difficulty with the VAS (Stucki *et al,* 1995). Difficulty with VAS has also been observed in older populations, where it has been demonstrated that NRS are preferred to VAS (e.g. Gagliese *et al,* 2005). Another important consideration with VAS is the anchor wording which can influence results (Streiner and Norman, 2008). This was demonstrated in a recent study of the DAS28 where the PtG VAS is not standardised. Five different PtG VAS wordings and anchors were completed by patients and DAS28

were calculated resulting in different DAS28 results that could be clinically significant in terms of access to anti-TNF therapy (French *et al,* 2013). In addition, practical limitations of VAS include difficulty in use over the telephone and distortion of line length when photocopied (McCormack 1988, Snow and Kirwan, 1988; Hawker *et al,* 2011), which may limit applicability in research and clinical contexts. Overall, despite the simplicity of VAS, it has been suggested that other approaches produce more accurate assessment and are preferred (Streiner and Norman, 2008).

6.3.2.2.2 NRS

NRS are similar to VAS but include numbers at regular intervals and can be 11, 21, or 101-point scales (Williamson and Hoggart, 2005). NRS have been suggested to have good completion rates by responders. In a recent systematic review looking at the assessment of pain intensity using NRS, VAS, and adjectival scales, 54 studies were identified (Hjermstad *et al,* 2011). Of the 19 studies that reported on compliance (including ability to complete, correct responses and error rate), most (n=15, 78.9%) reported better compliance using NRS. Overall, 11 studies recommended the use of NRS due to higher compliance rates, responsiveness, and ease of use (Hjermstad *et al,* 2011). However, this review did include articles reporting on broad populations including elderly and cognitively-impaired patients. It also reported that the articles reported compliance inconsistently across studies (Hjermstad *et al,* 2011). In another study comparing assessment using VAS and NRS in AS, the NRS was suggested as more effective as it took patients less time to complete (Akad *et al,* 2013). This is reinforced in an AS population where the assessment of spondyloarthritis international society recommend the use of NRS over VAS for BASFI and BASDAI (Sieper *et al,* 2009). Finally, from a practical perspective, NRSs benefit from the ability to be completed over the telephone.

6.3.2.2.3 Ordinal scales and Likert scales

Ordinal scales (also referred to as verbal rating scales or adjectival scales) provide descriptors along a continuum and are often used in self-reported health assessment (Streiner and Norman, 2008). In a cognitive interviewing study regarding difficulties completing questionnaires it was reported that patients prefer verbal statements to numbers (Meyer, 2007). A narrative review into different scales (VAS, NRS, and verbal rating scales), again in the context of pain intensity concluded that verbal rating scales were easy to use but not as sensitive as VAS or NRS, although it did highlight that the verbal rating scales format is under-researched (Williamson and Hoggart,

2005). Likert scales are comparable to adjectival scales however, rather than assessment being unipolar, Likert scales are bipolar and assess the full range of an attribute (e.g. strongly disagree to strongly agree) (Streiner and Norman, 2008). It has been suggested that Likert scales are common in nursing research (Rattray and Jones, 2007).

Other scales with which patients may be familiar were also considered. The response options of other PROMs employed within the rheumatology literature or that had been identified earlier within this study were explored. A number of PROMs including HAQ (Fries et al, 1980), Routine Assessment of Patient Index Data 3 (RAPID3, Pincus et al, 2008), and the Patient Activity Scale (PAS, Wolfe, Michaud and Pincus, 2005) use four response options based on difficulty (without any difficulty, with some difficulty, with much difficulty, unable to do). The BRAF also uses four response options, again based on difficulty but with slightly different wording (not at all, little, quite a bit, very much). Other scales such as the NHP (Hunt and McEwen, 1980) and the Rheumatoid Arthritis Quality of Life questionnaire (RAQoL) (Tijhuis *et al,* 2001) use binary (yes, no) response options. Finally, the recently developed Rheumatoid Arthritis Impact of Disease score (RAID) (Gossec *et al,* 2009; Gossec *et al,* 2011) uses individually worded 11-point NRS. Although not gathered in a comprehensive scoping review, this provided an idea of the response options currently in use in a number of regularly used scales within the rheumatology literature.

Finally, the patient perspective regarding response options was considered. During Study 2 there was some discussion regarding the patient perspective of response option formats. This included identification of advantages and disadvantages with VAS, and the importance of simplicity and low patient burden. However, no conclusive preference regarding response option format was generated. Despite this, data generated from discussion with patients did inform decision-making regarding appropriate response options.

As a result of consideration of the literature and qualitative work, items with different response options were developed. As recommended, formulation of response options should occur alongside item development to ensure compatibility (DeVellis, 2012). As a number of different ideas regarding item stem questions were being considered (Section 6.3.2.1) it was felt that different approaches may be required for different items and should be developed on an item-by-item basis. Given earlier discussion regarding the number of response options, it was decided that fewer response options

would fit with the literature and the patient perspective, and would be used where possible but the specific wording of response options would be developed alongside each item. During the development of the impact items which were generated directly from qualitative data it was felt that a standard response option could be used for all items in this section. During development, items were tested with different response options, particularly those identified from other PROMs used in rheumatology to identify response options that matched the question wording. The wording used in the BRAF-MDQ (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b) appeared to be the best fit for the developing items in this section. Furthermore, given the considerable work conducted in the development of the BRAF to ensure the patient perspective (Nicklin, 2009; Nicklin *et al,* 2010a; Nicklin *et al,* 2010b) this was felt to be a rigorous and appropriate option. The impact triad items have been assessed previously for fatigue using NRS (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b) but given the above discussion, both these and the response shift item were taken to Study 3 in different formats for further investigation into the most appropriate response option from the patient perspective. Response options for all other items were developed on an item-by-item basis.

6.3.2.3 Timeframe

It was also important to consider the timeframe over which items would be assessed. In PROMs or other questionnaires, respondents are required to answer questions within a specified time period which can be immediate (e.g. now) or more long-term (e.g. over a year). Bias can result from inappropriate recall periods (Stull *et al,* 2009) and the most appropriate recall period in the development of PROMs is an area of debate (Stull *et al,* 2009). In a review of the literature around recall periods it was identified that there are two broad categories of factors that influence responder recall; the characteristics of the concept of interest, and the context of the concept of interest to the responder (Stull *et al,* 2009). When defining the recall period for a PROM it is important to consider both the ability of the responder to provide accurate information within the specified timeframe and what is most appropriate for the purpose of the instrument (USDHHS FDA, 2009). Norquist *et al.* (2012) defined criteria for consideration in the selection of the length of the recall period for PROMs. These included consideration of the nature of the concept of interest (e.g. natural temporal fluctuation), the purpose for which the outcome measure is being used (e.g. intervention evaluation), the ability of the responder to recall the required information and the burden that poses (e.g. participant recall capacity), and finally the context in which the PROM will be used (e.g. study type). These criteria were used to address

the most appropriate recall period in the development of these preliminary stiffness items (Norquist *et al,* 2012).

In relation to the nature of the concept of interest, from the qualitative data generated in Studies 1 and 2 it was clear that the patient experience of stiffness was highly variable. Stiffness appeared to fluctuate with disease activity, and also vary within a 24-hour period. Variability was thus important to take into account when deciding on the recall of any stiffness item. In addition to variability, although stiffness was reported to occur in the morning period, the traditional concept of EMS/MS was challenged by patients. However, the current understanding of stiffness pathophysiology indicates a natural temporal fluctuation of stiffness in the early hours of the morning. These considerations were important to consider in decisions regarding timeframe.

In relation to the purpose of the outcome, it is the aim that the new stiffness tool could be used in clinical or research contexts. In terms of research, although there is no specific study for which this tool is being developed it is likely that any tool used in a research trial would be included as part of a questionnaire pack for example, for completion at baseline and at other time points during interventions. From this perspective, relative consistency of the timeframe with other PROMs may be useful, if it is clinically or biologically appropriate to the concept being measured. In terms of clinical use, the only consideration would be that the recall period provides useful enough information to inform decision making. In relation to the recall requirements and burden, the nature of the disease in this population will not the influence participant's ability to respond as might be the case in other chronic conditions such as dementia. However, this population is likely to be familiar with completion of PROMs given that many of its primary symptoms are patient reported (e.g. function, pain, and fatigue). The most frequently employed PROM used in trials involving RA patients is the HAQ (Kalyoncu *et al,* 2008) which asks responders to rate their functional ability over the past week (Pincus *et al,* 1983). In addition to these considerations, participants in Study 2 debated the advantages and limitations of different timeframes for items. However, no definitive timeframe was identified as optimal from the patient perspective. When discussed with the supervisory team, discussions centred on the purpose of use for the newly developed tool and highlighted that it was likely that in relation to interventions, the tool would need to be able to identify changes over days, weeks and months, but not hours (for example, it would be more likely to explore whether a course of timed release glucocorticoids

reduced stiffness in the subsequent days rather than whether analgesics relieve your stiffness within 30 minutes). From a clinical perspective it was felt that clinicians would expect some variation in a condition such as RA thus identifying sustained change would be more useful than short-term variation. Overall, it was felt that a short recall period (the past week or less) would be appropriate given the nature of stiffness. A very short recall period such as 'now' would only suit very frequent or variable concepts, and such approaches may also not capture relevant aspects of the patient experience (the purpose of the PROM) (Patrick *et al,* 20011b), therefore the past week was felt to be more appropriate. It was felt that the past week was an appropriate timeframe that would be acceptable to the population, consistent with other PROMs, and useful in terms of the information that it would provide. The past week could be used for any newly developed items and additionally, the traditional stiffness questions would capture different timeframes for comparison and exploration.

6.3.2.4 Layout and format

As formatting can improve response rates (Fanning, 2005) it was the final area to address during item development. There are a number of considerations within layout and format including reading level and style. It is generally recommended that completion of scales should only require a reading level of a 12-year old (Streiner and Norman, 2008). Similarly in health settings it has been suggested that reading levels required for health literature should be between 10 and 14-years old (Chapman and Langridge, 1997). There are a number of ways to assess reading levels including the Flesch Reading Ease Formula (Kincaid *et al,* 1975), Flesch-Kinkade grade level (Kincaid *et al,* 1975), the Gunning Fog Index (Gunning, 1968), and the Simple Measure of Gobbledygook (SMOG) (McLaughlin, 1969). These assessments generate scores based on the number of words per sentence and syllables per word. Assessment of the reading age of the newly developed items could be explored during development (prior to cognitive interviews) and involved testing of all preliminary items. Both the Flesch Reading Ease Formula and Flesch-Kinkade grade level (Kincaid *et al,* 1975) can be generated using Word, while the Gunning Fog Index and SMOG scores can be generated on the internet. Following their review of the literature, Chapman and Langridge (1997) identified a number of practical ways to improve the readability of health literature. Specifically in the development of PROMs, the use of short sentences (10-12 words) and avoidance of complicated words is recommended (Adams *et al,* 2013). The above recommendations were consistent with discussions with the supervisory team. Consideration of all recommendations

were used in the development of items during this study. The reading age of the preliminary items was assessed at the end of the development process (Table 6.2). Reading age requirements for the preliminary items met the recommended guidelines above (Streiner and Norman, 2008; Chapman and Langridge, 1997). They were also consistent with the reading age required (assessed by the Gunning Fog Index and SMOG) for other commonly used PROMs in a rheumatology context (Adams *et al,* 2013). The readability of the preliminary items was comparable to that of the HAQ (Fries *et al,* 1980) and the RAQoL (Tijhuis *et al,* 2001), reported in the study by Adams *et al.* (2013). However, as acknowledged by Adams *et al.* (2013) despite meeting readability recommendations, 22% of the UK population would still be unable to complete these PROMs, highlighting broader issues regarding their application.

Table 6.2: Reading age required for the preliminary stiffness items

The Gunning Fog Index	SMOG	Flesch Reading Ease	Flesch-Kinkade grade level
13 years	11 years	12 years	11 years
NB All 77 items were tested together to generate a score for each assessment method			

A number of suggestions were also made by patients in Study 2 regarding the layout and format of the questionnaire. Suggestions included preferences for simple marks rather than lots of writing, and that circles may be more difficult than a cross or a mark. The preference for simplicity was also reaffirmed, as was clear instruction describing the purpose of the questionnaire. These aspects, and those highlighted above (Chapman and Langridge, 1997) were taken into account prior to cognitive interviewing (Study 3) and were further discussed with the supervisory team following the outcome of item reduction and testing of the results of the final PROM (Study 4).

6.3.3 The final set of items for further testing

Following the process of item development involving iterative rounds of discussion with members of the supervisory team and consideration of the above elements, a set of 77 preliminary stiffness items was finalised to be taken forward for further testing in cognitive interviews (Study 3). The item tracking matrix (Appendix P) includes the final version of the 77 individual items (question numbers) within the 5 sections identified above, which were taken forward for further testing: 1) severity items (1-7); 2) impact items (8-59); 3) attribution items (60-68); 4) traditional stiffness items (69-74); and 5) response shift items (75-77). Chapter 7 will next describe the use of cognitive interviews to evaluate and further refine these preliminary items.

Chapter 7: Testing the draft content of the RA stiffness PROM with patients (Study 3)

Chapter 6 discussed the process of the development of draft items for an RA stiffness PROM. This chapter will discuss the process of testing and subsequently refining items with a relevant patient population.

7.1 Background

Previous Studies (1 and 2) used qualitative investigation to understand the patient experience of stiffness. These data enabled development of preliminary items for an RA stiffness PROM (Chapter 6). The next step was to address whether the items were acceptable to and understood by the target population. Potential problems with items include whether the wording is appropriate and whether the response options provided are suitable (Conrad and Blair, 1996; Drennan, 2003). Cognitive interviews are a method of critically evaluating products that provide information such as questionnaires, forms or brochures (Willis, 2005). They can also be used for item development for poorly understood concepts, questionnaire translation, and questionnaire development for populations where there may be difficulties with questionnaire completion (Drennan, 2003). Specifically, the use of cognitive interviews are recommended in PROM development guidelines (Patrick *et al,* 2011a; Patrick *et al,* 2011b) to ensure that that items are understood by patients in the intended way and difficulties with wording, response options, timeframe or format can be addressed. Cognitive interviews have been identified as a vital part of scale development as they are the last opportunity to adjust a measure prior to quantitative testing (Patrick *et al,* 2011b).

7.2 Aims and objectives

The aim of this study was to develop the draft content for a new RA stiffness PROM. It specifically aimed to test items with patients using cognitive interviews to ensure each item was clear, acceptable, and understood in the intended way, and to enable refinement prior to quantitative testing. The specific study objectives were:

- To test whether the instructions, items, and response options were clear and understandable to patients
- To investigate patient preference in relation to different item formats
- To investigate patient perspective in relation to different item wording

7.3 Methods

7.3.1 Cognitive theory

The cognitive aspects of survey methodology (CASM) approach has been influential in questionnaire design. The CASM approach identifies that "reporting errors in surveys arise from problems in the underlying cognitive processes through which respondents generate their answers to survey items" (Tourangeau, 2003, p.5). The cognitive processes required to respond to survey items have been captured in a four-stage cognitive model where the participant has to understand the item (Understanding); retrieve the relevant information from memory (Retrieval); make a decision about what information is relevant (Judgement); and match their response with the response categories provided (Response) (Tourangeau, 1984). To test the survey response process in relation to the cognitive model, cognitive interviewing is used (Willis, 2005).

7.3.2 Cognitive interviewing methodology

Cognitive interviewing is a method of testing survey instruments (Collins 2003; Beatty and Willis, 2007). 'Think aloud' and 'probing' are the two main cognitive interview techniques. During the 'think aloud' method, participants are instructed to 'think aloud' (i.e. articulate their interpretation and reasoning) as they read and respond to items. The interviewer has little input other than to encourage the participant to continue to speak (Willis, 1999). The alternative method of 'probing' allows the interviewer to ask questions regarding specific information following the participants' response to the item (Willis, 1999). This approach was developed for pragmatic reasons such as providing useful information for researchers (Beatty and Willis, 2007). There are benefits and limitations of each approach. In a summary of these approaches it was reported that 'think aloud' requires little interviewer training, may generate unexpected information and minimises interviewer bias, but is more burdensome for the participant partly because 'thinking aloud' can be difficult to do. 'Probing' on the other hand is easy for participants and enables the interviewer to maintain control of the session, but it allows the potential for interviewer bias (Willis, 2005). However, despite the apparently separate approaches, it has been suggested that researchers do not necessarily have to decide on the most appropriate method as for best results, aspects of both approaches should be integrated. Collins (2003) recommends that 'think aloud' and 'probing' methods can be combined and Willis (2005) suggests that in the practical application of cognitive interviewing "'think aloud' and 'probing' actually

fit together very naturally" (p.57), and as such researchers should adopt a flexible approach that incorporates aspects of both approaches for optimal results.

7.3.3 Participant identification and sampling

Participant identification and sampling were performed as described for Studies 1 and 2 (Section 4.3.2) but at the BRI only. A study specific PIS (Appendix Q) was used during recruitment.

7.3.3.1 Sample size and sampling

Often in qualitative research, little attention is given to sample size recommendations (Guest, Bunce and Johnson, 2006). This is also true for cognitive interviews where sample size recommendations have been highlighted as an area requiring further research (Beatty and Willis, 2007). Within the guidance available, 5-15 cognitive interviews have been recommended. Rounds of interviews of such numbers can be performed followed by review and interpretation (Willis, 2005). Given these recommendations, it was decided that the target for recruitment would be 10-15 participants. If further cognitive interviews were required because new problems were still being identified, recruitment could be continued.

7.3.4 Cognitive interview topic guide development

A topic guide (Appendix R) was developed to achieve the aims of the study (Section 7.2). The topic guide began with a set of instructions to introduce each participant to the cognitive interview procedure. This included emphasising the purpose of the interview and to reassure participants that they were in a safe environment in which they could speak freely. The set of instructions (Box 7.1) was adapted from the literature (Willis, 2005) and was discussed with each participant prior to commencing the interview. The preliminary PROM items that were developed in Chapter 6 (Appendix P) were formatted into a questionnaire pack for completion by participants during cognitive interviews (Appendix S), with input from the supervisory team and a patient partner (GB). A number of probes were also drafted for use by the interviewer throughout cognitive interviews (Appendix T). These probes were developed based on consideration of the questionnaire appraisal system (QAS) (Willis and Lessler, 1999). The QAS was originally developed for identifying sources of error when performing surveys over the telephone but can also be used for cognitive interviews (Willis, 2005). It is a checklist of items across eight categories of common error sources: 1) reading; 2) instructions; 3) clarity; 4) assumptions; 5) knowledge/memory;

6) sensitivity/bias; 7) response categories; 8) other (Willis and Lessler, 1999; Willis, 2005). The use of such a checklist provided an opportunity to consider and investigate suspected problems in items and enabled a systematic approach to probing (Willis, 2005). Therefore, the preliminary items were considered in relation to the categories included in the QAS which were used to develop a list of probes (Appendix T).

Box 7.1: Cognitive interview instructions for participants

Thank you for coming in. Let me tell you a little more about what we will be doing today:

- We are not collecting information about you but are trying our items out on people like you so we can make the items better
- Our goal here is to get a better idea of how the items are working. So I'd like you to 'think aloud' as you answer the items – just tell me *everything* that comes to mind as you go about answering them
- At times I might ask you about what you think an item is asking about, how you come up with your answers and how you interpret the items
- Some of the items might look very similar. This is because we are trying to find the best way to word the item, so if there are things you particularly like or don't like please do say!
- If any item is unclear, hard to answer or doesn't make sense please tell me that – don't be shy! There are no right or wrong answers.
- Finally, we will do this for about an hour unless I run out of things to ask you before then
- Do you have any questions before we start?

(Adapted from Willis, 2005)

7.3.5 Cognitive interviewing procedure

All cognitive interviews took place in non-clinical rooms in the Academic Rheumatology Unit at the BRI. They were performed by one researcher (Halls) who was unknown to participants prior to the study and who introduced herself as a doctoral student with a non-clinical background. Each participant was greeted by the researcher and was provided with refreshments. Prior to starting each interview, each participant gave informed consent and completed a questionnaire pack (Section 4.3.4.1). The instructions for participants (Box 7.1) were explained and each participant was asked if they had any questions prior to turning the audio recorder on. Each participant was asked to complete the draft items as they would any questionnaire but were asked to 'think aloud' as they did so. The researcher spoke if the participant stopped talking, and followed up silences, hesitations and questions from the participant with prompts or encouragement. Each cognitive interview was

audio-recorded and transcribed verbatim by a transcription service. All transcripts were checked for accuracy and anonymised by the researcher.

7.4 Analysis

As the process of analysis of cognitive interviews has been described as the major drawback in questionnaire pretesting due to the lack of standardisation (Drennan, 2003), the questionnaire testing literature has attempted to address this issue using taxonomies of possible problems (e.g. Conrad and Blair, 1996). These taxonomies are usually based on the four headings of the four-stage cognitive model: Understanding, Retrieval, Judgement and Response (Tourangeau, 1984; Collins, 2003; Drennan, 2003). As such, analysis in this study was deductive and coded under the four headings of the cognitive model. Coding was performed by the researcher and discussed with the supervisory team. The Nvivo 10 (QSR International Pty Ltd, 2010) software package and Microsoft Office Word and Excel 2013 were used for analysis.

7.5 Results

7.5.1 Participants

Of the 84 potential participants approached, 12 agreed to participate (12% recruitment rate). One participant who originally agreed to participate cancelled our interview at short notice and was unable to rearrange therefore, 11 RA patients participated. Seven were female (64%), age range between 51 and 83 years and disease duration between one and 25 years (Table 7.1). Reasons given for declining participation were as described in Chapter 4 (Section 4.3.4) however, the influence of multiple study recruitment will likely have affected participation.

Table 7.1: Participant demographic information

Pt ID	Gender	Age (yrs) †	Dis dur (yrs) ‡	HAQ §	PtG ¥	Pt pain ¤	Current medication	Work status	Education
2401	F	83	20	1.5	4.8	3.1	Analgesics, NSAIDs	Retired	College/apprenticeship
2402	F	61	2	1.375	5.9	4.7	Analgesics, DMARDs	RIB	School
2403	M	61	20	1.5	5.8	2.9	Analgesics, NSAIDs, Bios	RIB/Carer	School
2404	M	73	25	Inc.	6.1	7.1	DMARDs, Bios	Retired	School
2405	M	78	16	0.5	2.4	1.6	DMARDs	Retired	College/apprenticeship
2406	F	67	8	2	6.8	6.9	DMARDs, GCs	Retired	School
2407	F	51	1	1	7.3	9.5	Analgesics, NSAIDs, DMARDs, GCs	Paid work	University
2408	F	61	1	0	0.8	0.9	DMARDs, GCs	Paid work	School
2409	F	52	5	0.5	3.3	7.0	Analgesics, NSAIDs	Paid work	College/apprenticeship
2410	F	77	3	1.625	7.2	5.9	DMARDs	Retired	School
2411	M	56	3	0.625	4.3	5.0	DMARDs	Paid work	School
Median and interquartile range (IQR) †=61 (56-77), ‡=5 (2-20), §=1.1875 (0.5-1.5), ¥=5.8 (3.3-6.8) ¤=5.0 (2.9-7.0)									

Pt ID=Patient identification number; dis dur=disease duration; HAQ=Health assessment questionnaire 0-3 (3=most disabled); PtG=Disease activity score 0-10 (0=very well, 10=very badly); Pt pain=Pain assessment 0-100 (no pain-severe pain); NSAIDs=Non-steroidal anti-inflammatory drugs; DMARDs=Disease modifying anti-rheumatic drugs; GCs=Glucocorticoids; Bios=biologics; RIB=Receiving incapacity benefit; Inc.=Incomplete data

7.5.2 Cognitive interviews

All participants completed the draft items in the order specified in Appendix S, although the order of the three different formats (NRS, VAS and ordinal scale) of the impact triad items varied between interviews. In the first five interviews, all versions of the items were included in the questionnaire pack, with the NRS first, VAS second and ordinal scale last (as seen in Appendix S). Following comments by participants regarding the large number of items, in an attempt to reduce participant burden, the questionnaire pack was edited to contain only the one version of these items and the other formats were shown to participants for discussion only. Therefore, in the following three interviews the VAS format was included in the questionnaire pack and in the final three the ordinal scale format was included in the questionnaire pack.

Responses to each item are discussed below in relation to the headings of the four-stage cognitive model (Tourangeau, 1984). Items are presented in the five sections described in Chapter 6: severity items; impact items; attribution items; traditional stiffness items; response shift items. Each section provides an overview of the performance of items and participant quotes for illustration. Following analysis of the cognitive interviews, potential changes to items were identified and discussed with the supervisory team to generate the final wording. Changes made to items are reported below in boxes identifying the original and the refined item. If items were added, these were referred to as additional items. Changes to introductions and instructions are also discussed along with changes made for formatting and consistency reasons.

7.5.2.1 Severity items (1-7)

Generally items in this section were well responded to. Minor difficulties were identified under the headings of Understanding (n=4), Response (n=5) and Judgement (n=1). All items were refined for clarity, two items were added to enhance understanding and two were removed.

7.5.2.1.1 Have you experienced RA stiffness in your joints during the past 7 days? (item 1): Understanding and Response

This item was generally well understood by participants. However, two participants questioned the specificity of the word ‘joint’.

> *“Well, all joints that it’s affected by” [2404]*

"[…] when I get up it's a bit stiff to move my back […] now is that a joint?" [2405]

Two other participants questioned how many responses they could tick.

"So […] I can answer two boxes in one?" [2403]

"So I can just tick those can I?" [2409]

Minor changes to the wording, grammar and emphasis of the instructions were made for clarity (Box 7.2).

Box 7.2: Changes to item 1

Original item	Have you experienced RA stiffness **<u>in your joints</u>** during the past 7 days? Not in any of my joints ☐ Yes, in some of my joints ☐ Yes, in many of my joints ☐ Yes, in all of my joints ☐
Refined item	Have you experienced **RA stiffness** in your joints during the past 7 days? No, not in any of my joints ☐ Yes, in a few of my joints ☐ Yes, in many of my joints ☐ Yes, in all of my joints ☐

7.5.2.1.2 Have you experienced RA stiffness in your body (outside of your joints) during the past 7 days? (item 2): Understanding and Response

Item 2 was clearly understood by four participants who highlighted the relevance of this to their experience.

"Well yeah last week I just had it in […] all over stiffness. This one I'd say all over [item 2]. I just felt like I'd been in a boxing ring with a, you know, serious boxer, and just all over. So it wasn't just my joints I suppose is what I'd say there" [2407]

Six participants indicated uncertainty about this item.

"In your body, outside of your joints, oh heavens! Erm, when I've got a stiff neck, is that body rather than joints?" [2401]

"In your body, outside of your joints, that'll be a no I haven't had any in my body, what do you mean in your body?" [2402]

To understand these responses further the researcher explained what the item was trying to capture. Four participants felt that the concept was not relevant to their experience of stiffness. For some this was more relevant to aspects external to RA stiffness including ageing (n=1), overdoing physical activities (n=1), and pain (n=3).

"Oh yeah I think you do [get stiff everywhere] but I mean […] I am 60 odd, what do you expect?" [2402]

"Yes. I see what you mean, […] when I've had good days and I've gone in the garden and I've overdone it […] and I just feel I want to get in the bath […] because it does feel all over, tense muscles and things, so outside of your joints, I suppose you could say yes if you're talking about your muscles […] I don't know if RA affects the muscle" [2403]

"Yes, that could be pain as well" [2409]

This item was identified as not relevant to some participants' experience and two participants were unable to provide an appropriate response. However, other participants were able to respond as appropriate response options were available.

"Have you experienced RA stiffness in your body outside of your joints? I can't answer that, to be honest" [2403]

"Well I don't think I am bad enough to feel it except in bones. So not in any part of my body" [2401]

As a result of the difficulties highlighted by some participants, this item required changes to enhance clarity. As identified earlier (Chapter 4 and 5), stiffness 'all over' was not a concept that was relevant to everyone. However, removing the item might result in loss of information for participants to whom it was relevant. In later cognitive interviews, the researcher asked participants to whom the concept was relevant to suggest improvements to the wording.

"All over your body […] Yeah just ask them say have you ever experienced waking up and your whole body has gone stiff" [2408]

Following discussion, this item was divided into two items (Box 7.3). The first was similar to the original item with minor changes to grammar and emphasis, and wording that better reflected the patient experience and use of language (e.g. 'all over' rather than 'whole body'). The second item reflected the suggestion from the qualitative work

that all over stiffness was either present or absent. This was mirrored in the response options and wording, and was based on participant suggestions to better reflect the patient experience (Box 7.3).

Box 7.3: Changes to item 2

Original item	Have you experienced RA stiffness **<u>in your body</u>** (outside of your joints) during the past 7 days? Not in any parts of my body ☐ Yes, in some parts of my body ☐ Yes, in many parts of my body ☐ Yes, in my whole body ☐
Refined item	Have you experienced **RA stiffness** in your body (outside of your joints) over the past 7 days? No, not in any part of my body ☐ Yes, in a few parts of my body ☐ Yes, in many parts of my body ☐ Yes, all over my body ☐
Additional item	Over the past 7 days have you experienced **RA stiffness** all over? No ☐ Yes ☐

7.5.2.1.3 During the past 7 days have you experienced RA stiffness coming and going as frequently as usual for you? (item 3): Understanding and Response

Generally this item was understood by participants, however there was some uncertainty regarding the complexity of the item.

> *"[…] coming and going as frequently as usual, is that even English?" [2402]*

> *"Are they saying is coming and going within seven days, is that what it? […] so are they asking me do I normally have what I've had in the last week […]?" [2411]*

One participant suggested that the wording of the item assumes that stiffness is usual for everyone.

"Ah, you're making an assumption there [laughs] as usual for you. For me, it's unusual anyway because I've not really experienced any stiffness from RA since my elbows were replaced" [2405]

One participant also highlighted that the concept of usual as difficult due to variability.

"Well, I would say same as usual, which is random. [...] There is no usual." [2411]

During discussion, this item seemed to be asking a lot in one question, which might be contributing to the difficulties with Understanding and Response. When looking back at the qualitative data and early item development (Appendix O) the importance of both normality and frequency had been identified. Therefore this item was split into two items; one asking about normality (e.g. the usual experience of stiffness); and one asking about frequency (e.g. the occurrence of stiffness) (Box 7.4). Although the concept of stiffness being 'different to usual' had been important in earlier qualitative work and was reinforced by the patient partners, during cognitive interviews it was identified that this was not clear as a question. Therefore for clarity, the concept of 'usual' was retained in one item and the wording 'usual' was removed from subsequent items. Refinement of emphasis and wording was also performed.

Box 7.4: Changes to item 3

Original item	During the past 7 days have you experienced RA stiffness coming and going as frequently as usual for you? It has been **much less** frequent than usual ☐ It has been **less** frequent than usual ☐ It has been the **same** as usual ☐ It has been **more** frequent than usual ☐ It has been **much more** frequent than usual ☐
Refined item	Over the past 7 days has your **RA stiffness** been different to usual for you? It has been much better than usual ☐ It has been better than usual ☐ It has been the same as usual ☐ It has been worse than usual ☐ It has been much worse than usual ☐
Refined item	Over the past 7 days has your **RA stiffness** been as variable (coming and going) as usual for you? It has been much less variable than usual ☐ It has been less variable than usual ☐ It has been the same as usual ☐ It has been more variable than usual ☐ It has been much more variable than usual ☐

7.5.2.1.4 During the past 7 days have you experienced stiffness after a period of immobility (for example, in a chair or in bed)? (item 4): Understanding

This item was generally acceptable to participants. However, one participant suggested that the two tasks included in the example were not comparable.

> *"Bed, you assume, […] you've been in bed at night, which is a decent length, but […] if we watch a programme on television I can get up without using my arms or anything to get up from a chair, straight away afterwards, it doesn't [compare to getting up from bed]" [2405]*

As above (Section 7.5.2.1.3), the concept of stiffness being 'usual' was identified in the response options of this item.

"Again, you've got 'than usual'" [2405]

One participant highlighted that this item was not consistent with other items that specified RA stiffness.

"I think you need to separate those two really, the osteo and the rheumatoid" [2404]
"Do we need RA stiffness there?" [SH] "Yeah I think so yeah" [2404]

In response to participant comments, RA stiffness was added for consistency and the wording was changed to include only one action. To remove the emphasis on 'usual' and for consistency, the response options were changed to reflect increasing severity (Box 7.5).

Box 7.5: Changes to item 4

Original item	During the past 7 days have you experienced stiffness after a period of immobility (for example, in a chair or in bed)? I have had **much less** than usual ☐ I have had **less** than usual ☐ I have had the **same** as usual ☐ I have had **more** than usual ☐ I have had **much more** than usual ☐
Refined item	Over the past 7 days have you experienced **RA stiffness** after a period of immobility (for example, after sitting for a while)? No, not at all ☐ Yes, a little ☐ Yes, quite a lot ☐ Yes, very much ☐

7.5.2.1.5 During the past 7 days have your RA medications been controlling RA stiffness as usual for you? (item 5): Judgement

This item was identified by four participants as being difficult to answer as it required factual information.

"I don't know whether it's that that's stopping me from getting so much stiffness and pain" [2402]

"Oh, that's a difficult item […] it has been better controlled than usual the last 7 days because I've had the steroid injection, but the blood [still] shows that it's high" [2403]

This item had been included because it addressed an area that was discussed in detail during previous work (Chapter 4 and 5). However, given these participant comments it was removed as it appeared difficult for participants to answer and captured information that was contextual rather than useful for measurement.

7.5.2.1.6 During the past 7 days when have you experienced RA stiffness (in your joints or your body)? (items 6 and 7): Response

Items 6 and 7 both had the same wording but different response options (Box 7.6). Both items were generally well comprehended by participants. A number of suggestions were made in relation to the available response options. One participant suggested adding another response option in item 6.

"[…] you might want to put one in about, on getting out of bed first thing" [2405]

Two participants suggested additional response options for item 7 to provide comprehensive options.

"[…] first thing when I wake up, when I get out of bed, […] during the night when I get up it's bad, sometimes going to the bathroom, but does not when you are in bed" [2401]

Another participant highlighted a common difficulty with questionnaire completion based on how recently stiffness had been experienced.

"Can I just say that these items, I could answer them straightforward, when I've been going through it, but when you haven't got it, you forget you've had it" [2403]

The constant nature of stiffness was discussed by three participants. Some uncertainty was expressed about how to answer the item in relation to this however, generally it was felt that there were appropriate response options available.

"But my stiffness is like a continual stiffness […] so I don't know how to answer that one, morning, noon, or night, during the night […] I don't know what to put, to be honest, because it's there all the time" [2403]

"[…] it seems as though I am going to tick them all [laughs]" [2409]

Furthermore, one participant was unsure how many response options they could mark which indicated that clarification of the instructions was necessary. In discussion it was thought that it may be confusing and unnecessary to include both items. The response options provided in item 6 had greater simplicity and would limit different interpretations across participants. To improve item 6, a response option was added to capture if stiffness had not been present during the indicated times, and further minor changes were made to the instructions, emphasis and wording (Box 7.6).

Box 7.6: Changes to items 6 and 7

Original item	During the past 7 days **<u>when</u>** have you experienced RA stiffness (in your joints or your body)? During the morning ☐ During the afternoon ☐ During the evening ☐ During the night ☐
Original item	During the past 7 days **<u>when</u>** have you experienced RA stiffness (in your joints or your body)? First thing when I wake up ☐ When I get out of bed ☐ During the first few hours after I get up ☐ During the late morning ☐ During the early afternoon ☐ During the late afternoon ☐ During the evening ☐ During the night ☐
Refined item	Over the past 7 days when have you experienced **RA stiffness**? Please tick all that apply to you In the night ☐ In the morning ☐ In the afternoon ☐ In the evening ☐ None of these ☐

7.5.2.2 Impact items (8-59)

Items in this section were generally well responded to by participants. Each impact item (8-38) is discussed below, apart from items 8, 9, 12, 15, 21, 25, 29, 32, 34, 36 for which no difficulties were identified or no changes were made. Minor difficulties were identified under the headings of Understanding (n=18), Judgement (n=3) and Response (n=3). Five items were removed from this section (18, 26, 27, 33, 37, 38). All items in this section had consistent response options (not at all, a little, a lot, very much) which were not changed. The items developed to address the impact triad (39-59) are discussed separately (Section 7.5.2.2.21).

7.5.2.2.1 Has stiffness made it difficult to bath or shower? (item 10): Judgement

Four participants identified difficulties with the double-barrelled example provided in item 10. Some participants suggested that bathing was very difficult, others suggested that using a bath was not comparable to using a shower.

> *"[…] I am alright in the shower if I am stiff but not the bath" [2408]*

As a result the double-barrelled item was exchanged for wash, with an example provided. This broader wording was designed to ensure the item was relevant to a wider range of patients and the wording 'wash' had been suggested by participants in focus group 2 (Box 7.7).

Box 7.7: Changes to item 10

Original item	Has stiffness made it difficult to bath or shower?
Refined item	Has **RA stiffness** made it difficult to wash yourself (for example, have a shower)?

7.5.2.2.2 Has stiffness made it difficult to work? (item 11): Understanding

Five participants suggested that this item may be interpreted differently depending on participants' personal circumstance.

> *"Work, well of course I am not at work […] so work to me means keeping the house clean, cooking the meal, things like that" [2401]*

> *"When I think of work I'm just thinking about around the house, I'm my wife's carer […]" [2403]*

"Well, I'm retired [laughs] so […] what people regard as work is where they go out and earn some money […] but it's more than that, because we've done a load of voluntary work and raising money" [2405]

Following the first few interviews where participants indicated that work was broader than just paid work, the researcher asked participants in subsequent interviews whether combining item 11 (work) and item 12 (daily activities) would be acceptable. However, an important area of difference between the two items was noted by one participant.

"No, I think it should be different […] if you're a younger person and you've got to go to work, its different to being at home, retired and doing your chores. You can pick and choose when you do your chores […] you don't have to stick to the time factor or the routine of having to do it" [2406]

The researcher reviewed the earlier qualitative work to check patient quotes regarding the origin of this item. It was found that participant discussions in earlier qualitative work (Chapter 4) had related to daily activities, responsibilities and commitments including work, family life, making plans and childcare roles. However, as work was relevant to more participants, the researcher had taken that forward for item development without consideration of the other aspects. Therefore the item was edited with input from members of the supervisory team and checked with a patient research partner (GB) to attempt to capture all of these aspects (Box 7.8).

Box 7.8: Changes to item 11

Original item	Has stiffness made it difficult to work?
Refined item	Has **RA stiffness** made it difficult to carry out your responsibilities or commitments?

7.5.2.2.3 Has stiffness made it difficult to eat? For example, chew or cut your food? (item 13): Judgement

Three participants reported difficulties with the double-barrelled example provided in this item.

"[…] I think if you take that bit out there, chew. Cut food yes, it does, because I can't hold a knife and fork properly" [2404]

Again, the researcher referred back to earlier qualitative work to check the origin of this item. In Study 1, two participants had identified difficulties with eating and chewing as a result of jaw stiffness. During item development the researcher had then added

chew or cut as examples to make the item more broadly applicable. However, on consideration of this following participant comments it was identified that this was not appropriate. Therefore, the item was reworded to target the original aspect highlighted by participants (Box 7.9). It was also felt that cutting would be captured in other questionnaire items (e.g. item 21).

Box 7.9: Changes to item 13

Original item	Has stiffness made it difficult to eat? For example, chew or cut your food
Refined item	Has **RA stiffness** made it difficult chew?

7.5.2.2.4 Has stiffness made it difficult to do hobbies or activities you enjoy? (item 14): Understanding and Judgement

This item was generally well understood but some participants expressed concerns regarding its broad nature.

> *"You can certainly read however stiff you are […] but gardening is a hobby of course and that is not easy sometimes" [2401]*

In discussion it was felt that the broad nature of the item was important for relevance to as many people as possible, therefore only changes for consistency were made.

7.5.2.2.5 Has stiffness made it difficult to rise from a chair? (item 16): Understanding

This item was well responded to although three participants suggested their response would be dependent on the type of chair.

> *"[…] oh definitely to rise from a sofa, when you've got nothing to push on […] if it's got arms [its] much easier yes" [2401]*

> *"[…] sometimes can be, depending how low the chair is" [2409]*

The wording of this item was changed to focus on the action of getting up after being seated and increase emphasis on the action rather than the type of chair (Box 7.10).

Box 7.10: Changes to item 16

Original item	Has stiffness made it difficult to rise from a chair?
Refined item	Has **RA stiffness** made it difficult to get up after sitting for a while?

7.5.2.2.6 Have your daily activities required more effort than usual because of stiffness? (item 17): Understanding and Response

Generally participant comments regarding this item reinforced its relevance. However, one participant's interpretation of 'activities' indicated that the wording could be interpreted as only recreational activities.

> *"No I don't really do that much activities […] I used to do a little bit of dancing […]" [2409]*

This item wording was changed to include both tasks and activities and for consistency with earlier comments regarding the use of the term 'usual' (Section 7.5.2.1.4) (Box 7.11).

Box 7.11: Changes to item 17

Original item	Have your daily activities required more effort than usual because of stiffness?
Refined item	Have your daily tasks and activities required more effort because of **RA stiffness**?

7.5.2.2.7 Has stiffness had an impact on your daily life? (item 18)

This item was removed because the concept of impact was captured by other items, including another specific impact items (45, 52, 59), and to reduce participant burden.

7.5.2.2.8 Has stiffness made you slower? For example unable to rush (item 19): Understanding

Two participants identified different interpretations of the word 'rush' used in the example.

> *"No, that's not a good example for me. Has stiffness made you slower, yes, for example unable to rush. I do rush. Like I was saying earlier, this rush and this irritableness with the rush. I've got to get it done, things like that […] I know some people will say I am unable to rush, but whether the rush word is the right word. The item's okay 'til it gets to the rush" [2403]*

> *"Yeah. That is my problem, I rush around too much [laughs]" [2411]*

The word 'rush' was originally included as it was the wording used in earlier qualitative work (Chapter 4). However, on re-examination of the earlier data, being unable to perform things 'quickly' was also discussed. Therefore for clarity this item was reworded (Box 7.12).

Box 7.12: Changes to item 19

Original item	Has stiffness made you slower? For example unable to rush
Refined item	Has **RA stiffness** made you slower (for example, unable to do things quickly)?

7.5.2.2.9 Has stiffness made it difficult to do fine movements? For example, do up buttons on a shirt or cardigan? (item 20): Understanding

Some uncertainty was expressed regarding the examples provided in this item. One participant suggested that the example was seasonal and may not be relevant all year round while another suggested possible differences between genders.

> *"Well, it's this time of year, you don't do up buttons, do you?" [2403]*

> *"[…] quite a bit of female dressing involves hands behind the back [laughs]. Whereas us blokes we don't have that problem" [2405]*

It was felt that the example should be changed to something broadly relevant and gender neutral. On re-assessment of the earlier qualitative data the example of writing with a pen had been identified as difficult and was substituted as a more appropriate example (Box 7.13).

Box 7.13: Changes to item 20

Original item	Has stiffness made it difficult to do fine movements? For example, do up buttons on a shirt or cardigan?
Refined item	Has **RA stiffness** made it difficult to do fine movements (for example, write with a pen)?

7.5.2.2.10 Has stiffness made it difficult to make a fist? (item 22): Understanding

Three participants identified different interpretations of this item. One participant suggested that her experience would relate to both opening and closing her fist.

> *"Difficult to make a fist. No. Not now. But it was a few, you know a week or two ago, I couldn't get it, it went down like that [closed fist], well that's a fist alright but it would not straighten up [open fist]" [2401]*

Another participant suggested that the current wording sounded rather aggressive. While another suggested that this was not an essential everyday action.

> *"[…] you don't always make a fist, do you, you don't always need that to make a fist […] it's not like life and death, is it?" [2406]*

This item was originally developed because although many participants in previous qualitative work (Chapter 4) described difficulties with grip, some described difficulties specifically making a fist. Therefore both were included to identify which item performed better or whether they captured different information and were both important. To retain the original idea of this item, participant comments were taken into account and the wording was edited to include both opening and closing a fist. This aimed to make the item more broadly applicable and less aggressive sounding (Box 7.14).

Box 7.14: Changes to item 22

Original item	Has stiffness made it difficult to make a fist?
Refined item	Has **RA stiffness** made it difficult to open and close your fist?

7.5.2.2.11 Have you lacked physical strength to do your daily activities because of stiffness? (item 23): Understanding

Some participants identified the wording 'physical strength' as difficult to understand how this related to stiffness was also questioned. The earlier qualitative data were re-examined and 'strength' was identified as important during Study 2 (Chapter 5). However, the word 'physical' had not been used by participants, therefore this word was removed (Box 7.15).

Box 7.15: Changes to item 23

Original item	Have you lacked physical strength to do your daily activities because of stiffness?
Refined item	Has **RA stiffness** reduced your strength to do tasks?

7.5.2.2.12 Have you found that your movement is restricted because of stiffness? For example, reaching to get an item (item 24): Understanding

Although this item was generally acceptable to participants, difficulties were identified with the example provided. Two participants acknowledged that the example was quite specific and not relevant to their personal experience.

> *"[…] it depends where the stiffness is, doesn't it?" [2405]*

> *"Your movement is restricted, yes, not reaching out I don't think […]. Well you can reach because this part [upper arm] is usually okay on me but this part is worse [shoulder]. Do you see what I mean?" [2408]*

Following discussion and review of earlier qualitative data, the example was removed to ensure the item was as broadly relevant as possible (Box 7.16).

Box 7.16: Changes to item 24

Original item	Have you found that your movement is restricted because of stiffness? For example, reaching to get an item
Refined item	Has your movement been restricted because of **RA stiffness**?

7.5.2.2.13 Has stiffness made it difficult to move parts of our body or your whole body? (item 26): Understanding

This item generated some uncertainty regarding its wording and whether it was capturing similar information to item 24.

> *"No I don't think that's right, it's only my body, it does move, but it just doesn't move as smoothly as it should" [2407]*

During the original development of items 24 and 26, both came from discussions regarding movement difficulties. Both items were developed in an attempt to capture this idea and explore different approaches to asking about it. Given the similarity of these items, it was decided that item 26 would be removed to reduce the number of items and participant burden.

7.5.2.2.14 Has your body not moved like your brain tells it to because of stiffness? (item 27): Understanding

This item generated considerable discussion. The item resonated with some participants, although there were uncertainties about the wording.

> *"[…] I wonder if you could word that slightly different […] you want to do something, like pick up that thing, and your brain's telling you, you want to pick that up […] and you can't, because your hands or your shoulder, or whatever bit […] You want to do it and you want to reach it, but because of the restricted movement, your brain's telling you, yes you've got to move that mug, but it won't [go]" [2406]*

For these participants it was identified that this item captured similar information to item 28.

> *"This one is quite a good item. Well I don't know if you get the same from the two [27 and 28]" [2407]*

For other participants however, the item did not relate to their experience.

> *"Does your body not move like your brain tells it? No" [2402]*

Participant discussions in Studies 1 and 2 identified the cognitive impact of stiffness. Two items (27 and 28) were developed to capture this idea however, the item concept was quite abstract. Given the participant suggestion that these items captured similar information and following discussion, item 27 was removed to reduce participant burden.

7.5.2.2.15 Have you had to concentrate more than usual to move your body because of stiffness? (item 28): Understanding

Two participants discussed this item in detail. One participant suggested that the word 'concentrate' was only applicable to cognitive tasks while another indicated that this may be task dependent.

> *"I don't concentrate for that, it's just for reading" [2403]*

> *"[…] if you're doing fiddly little things you know, then yes, you'll have to concentrate more, but normal day to day things, no I don't think so […] If you're doing more intricate things I would say yes" [2404]*

Despite these comments, when this item was discussed, it was retained in favour of item 27. The item wording was edited slightly for consistency with other items (Section 7.5.2.1.4) (Box 7.17).

Box 7.17: Changes to item 28

Original item	Have you had to concentrate more than usual to move your body because of stiffness?
Refined item	Have you had to concentrate to move your body because of **RA stiffness**?

7.5.2.2.16 Have you felt worried because of stiffness? (item 30): Understanding

This item was generally considered acceptable although one participant identified uncertainty regarding the wording.

> *"Have you felt worried because of stiffness? Not particularly worried, concerned […] is this the beginning of something worse. But I don't go worrying about it for hours when there's nothing you can do at that time […] when they put worried in, it's concerned would be my word for that more" [2405]*

On review of earlier qualitative work (Chapter 5), 'worry' was the word that was used by participants. However, following discussion it was felt that by including both worry and concern the item may have broader relevance to participants (Box 7.18).

Box 7.18: Changes to item 30

Original item	Have you felt worried because of stiffness?
Refined item	Have you felt worried or concerned because of **RA stiffness**?

7.5.2.2.17 Have you felt embarrassed because of stiffness? (item 31): Understanding

This item was not identified as problematic for participants and although only one participant identified uncertainty regarding the wording, it generated some further discussion.

> *"No, it's awkwardness. Awkwardness, I think, and slowness, yes. No, I don't know embarrassed" [2403]*

Previous work on fatigue in RA had found that the word 'embarrassed' was difficult to translate (Nicklin *et al,* 2014). Synonyms included 'awkward', as suggested by one participant. However, the word 'awkward' was used in earlier qualitative work in general stiffness descriptions in the context of lacking ease of movement rather than in the context of embarrassment (e.g. *"[…] walking on stiff feet is just, it feels awkward […]" [102]*). Therefore 'awkward' was not felt to be an appropriate substitute. A

number of other options were discussed and the word ‘self-conscious’ was chosen as a substitute (Box 7.19).

Box 7.19: Changes to item 31

Original item	Have you felt embarrassed because of stiffness?
Refined item	Have you felt self-conscious because of **RA stiffness**?

7.5.2.2.18 Have you been unable to do your daily activities because of stiffness? (item 33): Understanding

This item was generally well understood by participants although it was identified that it captured very similar information to item 32.

> *“Being unable, didn’t it say that just now as well? To do your daily activities?” [2402]*

Another participant suggested that the item was too broad.

> *“That’s a bit all-encompassing isn’t it […] What, all of them? [Laughs]. It’s unlikely that you wouldn’t be able to do any at all” [2405]*

This item was removed as it was felt that item 32 captured similar information, as reflected in participant suggestion, and as no difficulties were identified with item 32 it was sensible to retain. On further consideration of item 33, it was recognised that the wording ‘unable’ was not suited to response options that gradually increased in severity (i.e. you are either able or unable, rather than a bit unable) providing further evidence for its removal.

7.5.2.2.19 Have you had to work around stiffness more than usual? (item 35): Understanding

This item was acceptable to participants although one was hesitant about its wording and asked for clarification.

> *“Work around, that mean like fathom out like ways of doing things? […] have you had to do things in a different way than usual because of your stiffness, that sounds better, doesn’t it, or is that me?” [2406]*

As identified earlier (Section 7.5.2.1.4), one participant was uncertain about the use of the word ‘usual’.

"Yeah. More than usual, but this is usual now […] what used to be usual is not what is usual now. You know, before I had arthritis" [2411]

During consideration of this item, it was felt important to retain the original patient wording ('work around') that led to the development of this item. However, to improve clarity an example was added based on participant suggestion. Additionally, the item was edited for consistency with other items regarding the use of the word 'usual' (Section 7.5.2.1.4) (Box 7.20).

Box 7.20: Changes to item 35

Original item	Have you had to work around stiffness more than usual?
Refined item	Have you had to work around your **RA stiffness** (or do things in a different way)?

7.5.2.2.20 Have you had to spend more time than usual coping (managing, dealing with, making do) with stiffness? (item 37) and Have you been able to cope (manage, deal, make do) with stiffness? (item 38): Understanding and Response

Items 37 and 38 were identified as capturing similar information.

"I don't think there's need for both" [2402]

It was also highlighted by two participants that the response options for item 38 were inconsistent with other items in this section.

"Your answers are going to be inverted here […] most of the 'not at all' answers it's not affecting you. Here you're saying you haven't been able to cope at all […] you're going to get a complete reversal. Here's me been going down, no, not at all, not at all and suddenly it's this, and have you been able to cope? Very much is the answer" [2405]

"Oh dear. Yes that's right. […] I think it's difficult because your head at this point isn't reading those [response options], so I didn't really read what that said […] This [Q38] is positive isn't it? Whereas this [Q37] is negative in a sense" [2407]

These items had originally been developed to capture direct self-management (see Appendix O). Different wording had been used given the uncertainty regarding the most appropriate wording for such items (Section 6.3.2.1.2.2). The identification of the reversed response options was not intentional and was very important. During discussion with the supervisory team it was felt that this information was captured in

other items within the impact triad (Section 7.5.2.2.21). Therefore both items 37 and 38 were removed to reduce the number of items and participant burden.

7.5.2.2.21 Impact triad items (39-59)

Items 39-59 were developed based on the concept of the impact triad (Sanderson *et al,* 2011). As previously discussed (Chapter 6), all items in this section were presented in three different formats (NRS, VAS, ordinal scale) and items developed to capture the self-management aspect were worded in four different ways. All items were discussed with participants to get a better understanding of patient preferences and identify the most appropriate format. To minimise repetition and discussion of items that were removed, this section will discuss participant comments in three respective sections; format, wording, and changes made to retained items.

7.5.2.2.21.1 Format

Discussion regarding preferences of item formats were broad and participants highlighted advantages and disadvantages of all formats. Generally, VAS were felt to be easy to complete and two participants suggested that they were accustomed to this format as it was often used as part of routine clinical assessment.

> *"No well that's [VAS] the same as what you do over the road [in clinic] isn't it? […] we are used to doing this" [2402]*

Another participant felt that VAS provided more flexibility than NRS and did not require much thought about the appropriate response.

> *"But I find this easier [VAS] because I don't have to think about it. I don't have to think about numbers, I can just put a mark on the line" [2407]*

One participant felt that VAS allowed greater honesty in response.

> *"I think sometimes you could be possibly more honest showing the definition on the line […] because it is like that you try and think oh well I won't moan or do the box above [on ordinal scale]" [2410]*

However, when discussing reasons for aversion to VAS formats, some participants felt VAS were imprecise, while others expressed preferences for NRS or ordinal scale formats.

> *"I don't like the lines no, I don't mind the circles" [2401]*

"[…] you could like […] put it in the wrong place, do you know what I mean, unless you read it […] whereas that is more clear [NRS], and those [ordinal scale]" [2406]

"I think if you grade it up to 10 it is probably easier in people's minds […] with the line I'd be thinking of going into colour coded, you know like the traffic light signal, extreme being red" [2409]

When comparing VAS to NRS, generally it was felt that NRS were easy to complete.

"I'm a numbers bloke […] but I still have difficulty in knowing where to put it [NRS], but I would, on this [VAS], even worse, I would say" [2405]

Two participants suggested that NRS or ordinal scale may be better for older people than VAS.

"I know where I would have to put it, but you get a lot of maybe older people than me, they would sooner see a number there" [2404]

"I think either of those [NRS or ordinal scale] is probably better than this [VAS], because like that could be quite confusing, especially if it's an older person" [2406]

Another participant suggested that VAS may be completed as if they were NRS.

"I know really there is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, even though they're not down there" [2403]

Ordinal scales were suggested to be easy to complete and to decide on an appropriate response.

"Yeah, actually I think boxes might be better for people […] because sometimes when I am over the road [in clinic] then I've got to fill one of these [VAS] out […] and I can't decide" [2402]

"I think these are quite easy to do like this […] I can't imagine too many people not being able to read and interpret that sort of thing" [2405]

Although ordinal scales were identified as being clear and easy to complete, the main complaint highlighted by three participants was the lack of options and flexibility in comparison to VAS or NRS.

"Because in a sense it doesn't give me a lot of room, does it, because it's that gap in the middle again […] It won't give me that option I'm looking for" [2403]

"These two [VAS and NRS] give you more flexibility to say a little bit, a lot. So for example, Q46 isn't quite as extreme as Q51, do you see what I mean, whereas if

I was answering those two on the boxes [ordinal scale], it would be the same box" [2407]

In contrast, one participant suggested that fewer options was an advantage.

"[…] some people might not like that [bigger range of options in the NRS], they might like that instead [ordinal scale], do you know what I mean, because you've still got the same answers" [2406]

Overall, given the limitations with VAS formats identified in the literature (Section 6.3.2.2.1), it would have required an overwhelming preference from patients to be considered. Given that this was not achieved, all VAS options were removed. The remaining ordinal scale and NRS formats were both supported by participant preferences in this study and Study 2 where participants indicated that fewer options placed less burden on the responder and also suggested that NRS were clear (Section 5.5.2.4.2). However, given the total number of items, it was felt that retaining items in both formats would increase participant burden and may influence completion rates. Therefore, the decision was made to retain only items in the NRS format.

7.5.2.2.21.2 Wording

When considering the items worded in different ways, the majority of participants indicated that these items were similar or that there was no need for all versions.

"No they mean the same. I think those really are tautology" [2401]

"They're basically saying the same, aren't they" [2411]

Only two participants defined these as distinctly different words.

"[…] so this item here about dealing with, you put a 6, and then this item here about coping you put a 5 […] what is the difference between these for you?" [Halls] "Well, its best how I cope, isn't it? […] It's not easy to deal with it but then after that you've got to cope with it. You've got to deal with it first before you can cope with it […] because the coping is the action, that's the way I see it" [2403]

"They are three different things […] We mainly deal with it, okay and we try to cope with it and some people find it harder to manage […] You can cope with it, but can you manage it?" [2408]

Others indicated slightly different definitions of different wordings but did not feel strongly that all three versions were necessary.

"Managing is more about what have you done about to help yourself, whereas coping is, for me, is more about how have you managed to get through" [2407]

"[…] everybody is going to deal, some give in more easier to some things than others don't they and some say well this is a bad day, I'll do what I got to do and then leave some things and I think you have just got to be sensible about it and try and do that haven't you really" [2410] "Yes, so the words to you, it doesn't matter which?" [Halls] "No, I mean I think you've put deal with it which I have been fairly, I have been honest about that, coping, yeah, and managing it" [2410]

With regard to participant preferences there was no outright favourite. Although two participants disliked 'made do'.

"No I don't think make do, it sounds a bit like 'make do and mend', you don't remember the war when you wanted a skirt a bit longer, you stick something round the edge of its bottom and you couldn't have any material because all clothes were rationed" [2401]

"[…] you make do, whatever […] that's how our parents grew up. They made do […]" [2403]

A further two participants indicated a preference for 'deal'.

"Which items would you have, if you had to design it?" [Halls] "Deal with […] and the others I wouldn't have" [2402]

"I think it would be just dealing with from day to day […] because it's all for the same thing really isn't it?" [2404]

However, across participants 'cope' or 'manage' appeared to be the most consistently preferred wording.

"Coping or managing. I think coping. Managing has sort of, slightly bossy lines but if you cope, you adapt to what you can do" [2401]

"[…] that's not bad, deal, but coped I think and managed, people would understand better" [2406]

"I think I prefer managed […]. Although all of them do, but managed is best I think" [2407]

In discussion and as highlighted earlier (Section 7.5.2.2.21.1), reducing the number of items was considered important. Therefore given patient suggestions and evidence from previous work where the wording 'cope' had been found to be effective (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b), this item was retained.

7.5.2.2.21.3 Changes made to retained items

Following the above considerations, only four items (39, 41, 44 and 45) from this section were retained for further testing. Of these, minor difficulties were identified with items 41, 44, and 45 under the headings of Understanding and Response. These difficulties and the subsequent changes made to these items have been detailed below.

7.5.2.2.21.3.1 Please circle the number that best describes how well you have coped with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days (item 41): Understanding and Response

There was some discussion regarding the placement of the anchors in this item. Some participants identified that the anchors were inconsistent with the other items. However, this was also acknowledged as positive with regard to improving the accuracy of responses as it would encourage participants to read each item fully.

> *"And this is the other end because 'very well' is at this end not this end this time" [2401] "[…] does that make more sense to be there?" [Halls] "It does to me because you've got all the goods this end and all the bad this end […] I think some people might find it a bit complicated […] On the other hand it might make them think more […] Perhaps that was why it was done" [2401]*

> *"That's quite interesting because those are the other way round as well aren't they? […] That's quite a difficult one because you are sort of answering the same thing, it's just that you are opposite" [2407]*

On discussion, it was felt that the anchor format for this item should be retained as it had been demonstrated to be effective in a validated scale (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b). However, in an attempt to reduce any influence of the anchor placement on response to other items, this item was placed at the end of the section. Minor changes were also made to the wording of all items in this section for consistency (Box 7.21).

7.5.2.2.21.3.2 Please circle the number that best describes the effect RA stiffness (in your joints or your body, and at any time of day) has had on your life during the past 7 days (item 44) and Please circle the number that best describes the overall impact on your life of RA stiffness (in your joints or your body, and at any time of day) during the past 7 days (item 45): Understanding

Three participants were uncertain about the wording of this item.

"[…] not everyone knows what effect is do they […] can't you put it more simpler like?" [2402] "I mean we can ask how important has RA stiffness been to you?" [Halls] "Yeah. I think that would be better than putting effect you know" [2402]

Others identified uncertainty regarding the difference between item 44 (importance) and item 45 (impact).

"My way of thinking it is quite similar" [2409]

Following discussion, the wording of item 44 was changed with 'effect' replaced by 'important' to enhance distinction between the items. Minor changes to all impact triad items were made for consistency and to enhance clarity (Box 7.21).

Box 7.21: Refined impact triad items (39, 41, 44, 45)

Please circle the number that best describes the impact that **RA stiffness** has had on your life over the past 7 days

No impact at all	0	1	2	3	4	5	6	7	8	9	10	A great deal of impact

Please circle the number that best describes the severity of your **RA stiffness** over the past 7 days

No stiffness	0	1	2	3	4	5	6	7	8	9	10	Extreme stiffness

Please circle the number that best describes how important **RA stiffness** has been in your life over the past 7 days

Not important at all	0	1	2	3	4	5	6	7	8	9	10	Very important

Please circle the number that best describes how well you have coped with your **RA stiffness** over the past 7 days

Not well at all	0	1	2	3	4	5	6	7	8	9	10	Very well

7.5.2.3 Attribution items (60-68)

Items 60-68 (Appendix S) were developed to explore the contribution of aspects identified as relevant to the patient experience of stiffness in Study 1 and 2 (Chapters 4 and 5). However, these items did not perform well during cognitive interviews. Although the ideas that each of the items attempted to capture were relevant to the

patient experience, they did not appear to be effective in a measurement format and participants identified a number of difficulties with them. Firstly these items were highlighted as being worded in a difficult way.

> *"I think 'contributed to your experience' is a little bit hard to get my mind round" [2401]*

> *"I am not quite understanding this item, I mean I do understand, I do but I don't you know […] I am picking it up but I am not picking it up as easily as I was everything else" [2402]*

One participant was uncertain about the word 'moderate' used in the response options. While others suggested the wording was unnecessarily repetitive making it complicated.

> *"I've read contribution five times and I am at two inches down the page" [2402]*

> *"The only thing I'm not so sure about though is moderate. What does moderate mean? […] moderate to me seems like oh only moderate, like in between" [2403]*

As discussed above (Section 7.5.2.1.5), participants identified difficulties with items that required factual information. This was highlighted in relation to the item regarding the influence of the weather.

> *"[…] well do we really know? It does sometimes seem it's the weather" [2401]*

This was also highlighted regarding the item about joint damage where participants were uncertain whether they had joint damage making it difficult to respond.

> *"I don't know what joint damage I've got, I don't know" [2402]*

> *"Joint damage. That one is a bit hard to know because unless you have an x-ray or something you don't really necessarily know, unless it is so visually obvious" [2409]*

Uncertainty was also highlighted in relation to making factual statements about the effect of medications.

> *"I wouldn't want to go into [Consultant in clinic] now and say my RA medications aren't controlling my symptoms, my disease because I don't think we know yet" [2407]*

Participants found the inexplicit nature of item 65 difficult.

"What other RA symptoms are there?" [2401] "What would other RA symptoms include for you?" [Halls] "Ha ha good answer, apart from stiffness and pain, I can't think of any" [2401]

"Other, well, does that include pain?" [2411]

Finally, some participants felt that if items were not currently relevant to them, there was no appropriate response option.

"[…] this is a horrid one in a way to answer because I haven't been in an RA flare so I suppose no contribution is what one would put" [2401]

"No, see this bit don't apply because I am not actually taking the medication that they required me to" [2409]

Given the numerous difficulties with these items, considerable thought was given to their inclusion. It was decided that although these items helped understand the context of the patient experience of stiffness they did not provide information that would be useful for measurement because they were difficult to answer. Given these considerations and other concerns regarding the length of the overall questionnaire they were removed.

7.5.2.4 Traditional stiffness items (69-74)

The traditional stiffness items from the literature were included for cognitive interviewing as cognitive debrief did not appear to have been performed before. Although this process was not intended to result in changes to items as it was important to retain their traditional wording and format, it was felt that information generated here may provide information to aid decision making regarding the subsequent retention or removal of items in Study 4. However, given the large volume of items taken to cognitive interviews and that this was not the key objective of this study, less time was spent exploring these items than was anticipated and less data were generated regarding these items than other sections in the questionnaire pack. Despite this, a number of participant comments relating to the headings of Understanding and Response are described below.

7.5.2.4.1 How would you describe the overall level of morning stiffness you have had from the time you wake up? (item 69 (VAS), item 70 (NRS), item 71(ordinal scale)): Understanding and Response

Some uncertainty was reported regarding the wording of 'overall level of morning stiffness'.

"How would you describe the overall morning, overall? Overall stiffness?" [2411]

One participant appeared to relate the item to duration rather than severity.

"Because how would you describe the overall level? […] because there's no like hours there. Although down here [item 73] there is" [2409] "So would something like that be easier to do you think?" [Halls] "I think so, yes" [2409]

Despite these considerations, no further difficulties were identified. For consistency with earlier discussion (Section 7.5.2.2.21.1), the item using a VAS format was removed. This was reinforced as appropriate by participant comments indicating that items with NRS and ordinal scale formats were easy to answer.

7.5.2.4.2 How long does your morning stiffness last from waking until maximum improvement occurs? (item 72 (minutes and hours), item 74 (ordinal scale)): Understanding and Response

Participants disclosed some uncertainty regarding response to this item, particularly in relation to item 72. One participant suggested this format was acceptable but imprecise and could be improved by removing the specified units of time.

"Yeah, I'm not sure whether you'll be able to put the exact amount down […] I think you could just put down there how long, I don't think minutes or hours is relevant really" [2402]

Another response suggested that the format was not clear enough that participants had to specify the amount of each unit.

"Minutes, hours, that could be hours, okay, not minutes" [2408]

Another participant highlighted that it was difficult to respond to this item for individuals with no stiffness because the item does not inherently provide a 'none' option.

"[…] you're leading you see, how long does your morning stiffness last?" [2405]

As in earlier qualitative work, some participants queried the focus on morning stiffness, highlighting the broader nature of the experience of stiffness.

"[…] it can last more or less all the day, but it usually eases up within about four hours" [2411]

Other participants were uncertain about 'maximum improvement'.

> *"Maximum improvement, well, what's maximum improvement?" [2411]*

In contrast, for some participants no difficulties were reported.

> *"[…] it's marred, you know, it gets marred or less. What I do, I got my own way of dealing with it. I just sit in the chair in the morning, and like come to yourself […]" [2406]*

7.5.2.4.3 Were your joints stiff when you woke up today? If yes, how long did this stiffness last? (item 73)

This item was presented in one format only and appeared acceptable to participants.

> *"About an hour" [2403] "Are those easy to answer?" [Halls] "Yes, because […] I'm sort of bringing it all together quickly and […] [giving the] first answer that comes into your head" [2403]*

Although large amounts of data were not collected in relation to traditional items, these data combined with data collected in Study 2 will inform decision making in Study 4. All the traditional items were taken forward for inclusion in Study 4 apart from item 69 which was replaced by another severity item. The item was taken from the preliminary flare questionnaire (PFQ) and has been demonstrated to be able to distinguish patients in flare from those not in flare (Bykerk *et al,* 2012; Bykerk *et al,* 2014b). This item was not identified in the systematic literature review (Section 2.4.1) but was highlighted during discussion at OMERACT. Although not cognitively debriefed with patients it appears to be the only traditional stiffness item which focuses on 'stiffness' rather than EMS/MS, which fits with patient descriptions from Studies 1 and 2. Given this, and that the item was clearly defined it was felt relevant to include as another traditional severity item to test in Study 4.

7.5.2.5 Response shift items (75-77)

Items 75-77 were developed in an attempt to explore the response shift. Like the impact triad items (Section 7.5.2.2.21) these items shared wording but were tested in three formats (NRS, VAS, ordinal scale). Generally these items were well responded to by participants. Minor difficulties under the headings of Understanding and Response were discussed. Firstly there was uncertainty regarding a 'usual week' and the ability to identify flare.

"Usual week, is there with arthritis anything such as a usual week?" [2401]

"I mean I am not used to identifying the flares really" [2401]

Another participant answered the item based on the anchors being reversed but this was clarified through discussion. Although identified as being potentially difficult, the item was retained but the wording was made simpler and only the NRS format was retained (Box 7.22).

Box 7.22: Changes to response shift items

Original:
Thinking about a **usual week** when you are **not in a flare** (flare-up) of your RA please circle the number that shows your **usual** RA stiffness (in your joints or your body, and at any time of day)

No stiffness	0	1	2	3	4	5	6	7	8	9	10	Extreme stiffness

Refined:
Please circle the number that best describes the severity of your RA stiffness over a usual week when you are not in a flare?

No stiffness	0	1	2	3	4	5	6	7	8	9	10	Extreme stiffness

7.5.2.6 Instructions, formatting and consistency

In addition to discussion regarding specific items, it was clear that other aspects including wording and format also required clarification. This included the introduction, and general formatting and consistency aspects where changes were made to enhance clarity and ease of completion. These have been described below.

7.5.2.6.1 Questionnaire introduction: Understanding and Judgement

There were some difficulties regarding understanding the introduction. One participant was uncertain about the term joint damage due to its factual nature, as highlighted previously (Sections 7.5.2.1.5 and 7.5.2.3).

"Yes, the only thing what throws me a little bit, it says, due to joint damage. Well […] have I got joint damage or has it just grown?" [2403]

Another three participants were uncertain about the instructions being given in the introduction and therefore what to include in their answer.

"I could do the questionnaire […] but am I really clear in my mind […] what that's asking? […] Because I wouldn't know what to put, to be honest, now I've read that bit in brackets especially" [2403]

"Don't include the fact that I can't straighten my arm or anything like that then?" [2405]

Four participants questioned whether they needed to write anything in the space underneath the introduction. One participant had to clarify 'RA' and another questioned whether this questionnaire took pain into account.

"[…] what's RA stiffness? Oh rheumatoid arthritis! [Laughs]" [2404]

"[…] do you want to know about [...] the difficulties in movement, regardless of pain?" [2411]

It was clear that the questionnaire introduction required changes to enhance clarity. Refined wording was generated through discussion with the supervisory team, in particular with one patient partner (GB) (Box 7.23). Minor formatting changes, including the removal of the space below the introduction and the use of emphasis were made to enhance clarity.

Box 7.23: Changes to questionnaire introduction

Original item	This questionnaire is about stiffness related to your rheumatoid arthritis or **RA stiffness**. It will help us understand **how active your disease is**. Some people have joints that are always difficult to move whether their RA is good or bad (for example, due to joint damage). Please do not include this sort of stiffness when you answer this questionnaire. We would like to know how **RA stiffness** has affected you during the **past 7 days**
Refined item	This questionnaire is about RA stiffness that comes and goes. It is not about joints that are permanently stuck (for example, due to an operation). However, we do appreciate that sometimes even permanently stuck joints do get stiffer (for example, when your disease is bad). Please just try to think about the stiffness that comes and goes as you answer this questionnaire.

In addition to the changes to the introduction, suggestions from the team indicated that it would be important to explore the relevance and influence of joint damage. Therefore two items to capture its presence and how it is reported by patients were developed (Box 7.24). The first additional item was developed to capture the presence of joint damage and was placed at the beginning of the questionnaire. The second additional item was developed to capture the amount of reported stiffness that was a result of joint damage and was placed at the end of the questionnaire. Item

wording was developed from patient descriptions from Studies 1 and 2 and the supervisory team including considerable input from one patient partner (GB).

Box 7.24: Additional items related to the introduction

Additional item	Do you have any joints that are permanently stuck? No ☐ Yes ☐
Additional item	How much of the stiffness you have reported in the items above is about joints that are permanently stuck? None of the stiffness I have reported ☐ A little of the stiffness I have reported ☐ Quite a lot of the stiffness I have reported ☐ All of the stiffness I have reported ☐

7.5.2.6.2 General instructions, format and consistency

A number of changes were made to the overall questionnaire based on comments from cognitive interview participants and discussion with the supervisory team. Firstly, in relation to the layout of the questionnaire, in some instances participants found that they had marked the wrong response option. Consequently the layout was edited to include boxes for clarity and ease of response.

> *"Oh, did you say that one was very much so?" [Halls] […] "Oh yes, got it in the wrong one, haven't I?" [2406]*

> *"I think now that I'm getting down to this level here, I think you probably could do with a line down here as well" [2407]*

Secondly, as identified earlier (Studies 1 and 2), participants highlighted relationships between symptoms such as stiffness and pain, and between different conditions such as OA and RA.

> *"I am not too sure if the stiffness affects my sleep, it is more if I am in a bit of pain but like I said again it is the combination of the two […]. Is it the pain or is it because I am stiff?" [2409]*

In discussion about these relationships it was felt important to reinforce the topic of stiffness in the context of RA throughout the questionnaire. 'RA stiffness' was incorporated into each stem question to reinforce the topic throughout the

questionnaire. Thirdly, it was identified that some items contained the word ‘over’ while others used the word ‘during’ to describe the timeframe. For consistency, ‘over’ was decided upon and used throughout the questionnaire. Finally, other minor changes were made for consistency and clarity. These included replacing any double reinforcement with single reinforcement using bold only, and clarifying and ensuring appropriate placement of all instructions

7.5.2.7 Final set of draft items

During the process of cognitive interviews and subsequent review with the supervisory team, 36 items (5, 7, 18, 26, 27, 33, 37, 38, 40, 42, 43, 46-68, 76 and 77) were removed and four items were added. The final 45 draft items included 39 new items and six traditional stiffness items which were taken for further testing in Study 4.

7.6 Discussion

Overall, the results from the cognitive interviews indicated that the draft RA stiffness PROM items were acceptable and understandable to patients with only minor difficulties identified. This demonstrates a key strength of the development of these items which involved a rigorous process of qualitative investigation that was informed at all stages by patients. The majority of difficulties identified fell under the heading of Understanding and mainly related to the identification of minor but necessary improvements regarding the wording and clarity of items. Although identified difficulties were minor, this study demonstrates the importance of using cognitive interviews to enhance understanding of PROM items in the intended population and in PROM development.

Although cognitive interviewing will not identify all problems with survey items, it is considered that the most significant problems will be highlighted by their use (Beatty and Willis, 2007). If cognitive interviewing had not been performed in this study, difficulties would not have been detected which may have led to inaccuracies in future data collection. The use of cognitive interviewing is also consistent with recommendations in the development of PROM instruments, particularly in relation to the content validity (Patrick *et al,* 2011b). The rigorous tracking of the development of items within previous Chapters (4-6) was especially useful where problems were identified with items. The transparent item development process enabled the researcher to return to earlier qualitative data to check ideas and wording to ensure

that the relevant concept was being captured or that appropriate patient language was included, again enhancing content validity. In addition, data generated from this study has provided further evidence directly from patients, regarding the format and wording of draft items. These results add to data generated from Study 2 and provide evidence to support the current literature (e.g. Section 6.3.2.2). Cognitive interviews also provided the opportunity to further explore patient preferences regarding the use of different item formats and wording. This was important to enable informed decision making regarding the tradeoff between rewording and removing items. An interesting observation from the cognitive interview process was the detection of no difficulties under the analysis heading of Retrieval. This is consistent with other recent work (Murtagh, Addington-Hall and Higginson, 2007; Nicklin *et al,* 2010a) and it has been suggested that the short timeframe (past week) employed in all these studies may explain this finding (Murtagh, Addington-Hall and Higginson, 2007; Nicklin *et al,* 2010a).

Taking forward only the items in an NRS format was a key decision. This was advantageous from the perspectives of practicality (Akad *et al,* 2013) and given the suggested improved psychometric properties compared with VAS (van Tubergen *et al,* 2002; Franchignoni *et al,* 2014). It will also reduce participant burden in Study 4. However, the use of an NRS scale is inconsistent with the optimal number of response options recommended by the seven (plus or minus two) rule (Miller, 1956, reprinted in 1994; Streiner and Norman, 2008) where it is suggested that participants are unable to discriminate responses over this recommendation. Although it was decided that only items in NRS format would be taken forward for further testing, both ordinal scale and NRS formats appeared to be acceptable to patients. Some scales such as the WOMAC are validated in multiple formats (Bellamy, 2005). Further investigation regarding item format could be explored in future research.

Despite the advantages of cognitive interviewing, there are limitations of the approach. The value of information generated from cognitive interviews has been questioned given that they involve artificial environments and small sample sizes (Drennan, 2003). The artificial environment created by cognitive interviewing may result in differences in the completion of items once the researcher is not present. For example, each cognitive interviews lasted approximately one hour but it is unlikely that participants would spend that amount of time completing items in applied environments. Despite this, this study was just one part of an extensive PROM development process involving earlier qualitative (Chapters 4 and 5) and subsequent

quantitative (Chapters 8 and 9) studies. Therefore, rather than producing entirely independent evidence, they instead provide evidence as part of a broader body of work, and cognitive interviews are most effective when used in combination with other validity and reliability assessment methods (Drennan, 2003).

In relation to sample size, it is acknowledged that the information generated from cognitive interviews is qualitative rather than quantitative in nature (Drennan, 2003; Willis, 2005), and therefore inherently associated with small samples. As expected in cognitive interview studies (Willis, 2005), the participant sample in this study was small (n=11), although it did include participants with a range of age, gender and disease duration. Also acknowledging the qualitative nature of cognitive interviewing data, the principal of data saturation (Guest, Bunce and Johnson, 2006) was used. Data saturation was felt to have been reached in this study although given the deductive analysis approach this was based on the generation of no new difficulties rather than no new themes.

Other limitations also relate to participant demographics. The 'think aloud' process has been suggested to be difficult for participants (von Thurn and Moore, 1994), particularly for certain groups of individuals such as those with low educational levels (Wellens, 1994). Although participants in this study did not appear to find the process difficult, participants were all required to be able to speak English unaided to be eligible to take part, and all participants indicated that they had at least a school level of education (Table 7.1). Therefore, the draft items have currently not been cognitively tested in individuals with self-defined lower levels of education or English language ability. However, attempts were made previously to ensure that the items were as understandable and accessible as possible, and met recommendations regarding readability (Section 6.3.2.4). Although, as acknowledged in Chapter 6, readability recommendations themselves have inherent limitations (e.g. Adams *et al,* 2013). The study sample is also limited by the lack of collection of information relating to ethnicity. Although no formal information was captured, all participants for all studies were recruited from hospitals in South West England and were Caucasian. Despite this, local colloquialisms were avoided during item development. Furthermore it would be expected that the conceptual underpinning of items would be relevant in broader populations. As discussed in Chapter 4, conceptual consistencies have been identified between this work and other research targeting the same topic, which also involved an ethnically heterogeneous participant sample (Orbai *et al.* 2014). Therefore data generated in this and preceding qualitative studies provide a good

basis for further item development and testing. However, further development of any new PROM may include translation and cultural adaption (USDHHS FDA, 2009).

A key strength of this study was the discussion of results with members of the supervisory team allowing enhanced interpretation, from a range of perspectives. Identifying problems with multiple reviewers is considered good practice and has been suggested to combat some limitations of cognitive interview such as ambiguity of participant responses (Conrad and Blair, 2009). Although the items were not cognitively debriefed in their final format, detailed discussion with the team's patient partners (GB and AE) enabled development and review of items from expert patient perspectives. Finally the use of the COREQ reporting framework enhanced the transparency of the study (Tong, Sainsbury and Craig, 2007) (Appendix U).

7.7 Conclusions

This study has allowed testing and refinement of the draft RA stiffness PROM items. Although predominantly the changes made to these items were minor, they were crucial in ensuring that they were understandable and acceptable to the intended population. This study has demonstrated the benefits of cognitive interviewing to reduce reporting errors during the completion of questionnaires. The 45 draft PROM items are now suitable for quantitative testing and validation (Study 4).

Chapter 8: Developing the structure and content for an RA stiffness PROM (Study 4, part 1)

Preceding chapters have described the process of development and testing of draft items for a new RA stiffness PROM using qualitative approaches to enhance content validity, and resulting in 45 draft items (39 new and six traditional). This chapter is the first of two describing quantitative methods to develop the most effective item structure for a novel RA stiffness PROM. It reports data collection and demographic description, then focuses on the identification of the most appropriate analysis to use for these data, describing the theoretical underpinning and presenting a worked example comparing two analytical approaches.

8.1 Background

8.1.1 Questionnaire development methodology

The importance of combining qualitative and quantitative methods in the development of PROMs has been emphasised in the literature (Patrick *et al,* 2011b). Consistent with recommendations (USDHHS FDA, 2009), previous qualitative studies were followed by a quantitative study to test the draft items. Surveys provide an opportunity to systematically collect information from large samples (Groves *et al,* 2009). Therefore in this study, a survey enabled collection of responses to the draft items from a sample of RA patients. The development of a new PROM was based on CTT. CTT is grounded on the idea that a participant's observed score is the result of true score plus error (DeVellis, 2012) and has traditionally dominated the field of scale development (Streiner and Norman, 2008). The theory of CTT is broadly applicable in many testing situations because its assumptions are considered easy to meet, but weak (not stringent) (Hambleton and Jones, 1993), which highlights some of the advantages and limitations of this approach. The newer, IRT framework broadly aims to overcome the limitations of CTT, and includes models such as Rasch (Streiner and Norman, 2008). The implications of using CTT are discussed in Section 8.5. The development of PROM using CTT requires a series of statistical analyses including initial assessment of the suitability of items for inclusion and exploration of whether the PROM contains different groups of items. This is then followed by analyses to identify the smallest combination of items that work well together to effectively evaluate stiffness (item reduction) (Pett, Lackey and Sullivan, 2003). This chapter describes data collection and demographic description, followed by a comparison of two analytical approaches to establish the most appropriate method which will then be used to develop and test the RA stiffness PROM (Chapter 9).

8.1.2 Multivariate analysis

The family of factor analysis techniques are often used in scale development to identify groups of related items (Pallant, 2010). The primary functions of what is broadly referred to as factor analysis are 1) understanding the relationships in a set of items; 2) development of a structure to assess the concept of interest; 3) reducing the number of items by retaining only necessary items without losing information (Field, 2009). There are different approaches to factor analysis including confirmatory factor analysis (CFA) and exploratory factor analysis (EFA). CFA is useful for hypothesis testing, when looking to test how well an *a priori* model fits the data (Pett, Lackey and Sullivan, 2003; Field, 2009). EFA is better suited to exploring data, when looking to understand the most appropriate factor structure from the data (Pett, Lackey and Sullivan, 2003; Field, 2009). As this study was exploratory in nature, it adopted an EFA approach. Within EFA there are a number of different techniques including principal component analysis (PCA), principal axis factoring, and image factoring (Field, 2009). As PCA is a straightforward, well recognised and commonly employed approach to factor analysis (Pett, Lackey and Sullivan, 2003) it was used here. However, it is important to acknowledge that although PCA falls under the umbrella term factor analysis, it has a different underpinning mathematical method (Field, 2009). There is considerable debate within the methodological literature about the strengths and limitations of the two approaches, where some do not even perceive PCA to be a member of the factor analysis family (Schmitt, 2011). However, the method of PCA is robust and less complex than factor analysis (Field, 2009) enabling better understanding in applied rather than methodological contexts. Furthermore, the two approaches have been found to produce similar results (Stevens, 2002). With regards to terminology, although PCA was the approach employed in this study, much of the literature refers to factor analysis to describe the many approaches within the family of factor analysis techniques. Therefore, subsequent references to factor analysis refer to PCA (under the broad umbrella of factor analysis).

8.1.3 The appropriateness of PCA

PCA has many advantages including being straightforward, well recognised, common, accessible and stable (Pett, Lackey and Sullivan, 2003; Field, 2009; Linting and van der Kooij, 2012). However, there are considerations regarding its appropriateness. Statistical tests require certain assumptions to be met to produce accurate results (Field, 2009). Generally, the assumptions of parametric tests are appropriate for PCA (Pett, Lackey and Sullivan, 2003) which assumes that there are linear relationships between variables (items), and that variables are scaled at an

interval or ratio measurement level (Linting and van der Kooij, 2012). As the variables to be assessed in Study 4 were collected on a combination of scales (dichotomous, 4-point ordinal scale, 5-point Likert scale, and 11-point NRS), it is unlikely that these assumptions would be met. Although variables assessed using Likert scales of five categories and more are often treated as continuous (e.g. Bollen and Barb, 1981), implications of the use of PCA with inappropriate data have been documented in the literature. For example, the use of Pearson's correlation coefficients has been questioned (Choi, Peters and Mueller, 2010) as they can underestimate correlations (Olsson, 1979) and lead to inaccurate factor loadings (Bernstein and Teng, 1989). Given these considerations, different approaches for performing factor analysis with non-continuous data have been recommended. One recommendation is the use of polychoric correlation coefficients for variables with ordered categories and tetrachoric correlation coefficients for dichotomous variables, to create the correlation matrix (Streiner and Norman, 2008; Field, 2009). The use of tetrachoric correlations were clearly inappropriate for these data as only two of the 45 draft items were dichotomous. Polychoric correlations can be viewed as a transformation that stretches the response scale to produce "corrected" correlations (Lorenzo-Seva and Ferrando, 2014, p.884). This makes polychoric correlations more appropriate for dichotomous and ordinal data (Streiner and Norman, 2008). However, there are limitations with polychoric correlations including that they are suggested to be less stable than Pearson's correlation coefficients because they are generated from a model-based estimate rather than generated directly (Chen and Choi, 2009). Furthermore, the option to produce these correlation coefficients is not provided in many traditional software packages (Baglin, 2014). Programmes that are available to generate polychoric correlations are often very basic and are unable to deal with missing data (BayesPCC, Choi, Chen and Kim, 2009; FACTOR, Lorenzo-Seva and Ferrando, 2013), and importantly expect all variables to have the same number of response categories (i.e. be assessed on the same scales) (Lorenzo-Seva and Ferrando, 2013).

Another recommendation for use with ordinal data is nonlinear principal component analysis (NLPCA) (Linting and van der Kooij, 2012). Consistent with the aims of PCA, the aim of NLPCA is to understand the structure of and reduce the number of items in a data set (Linting and van der Kooji, 2012). The strengths of NLPCA include that the approach can take into account non-linear relationships between variables and can include items with different levels of measurement (Linting and van der Kooji, 2012). A key limitation of NLPCA is that, similar to polychoric correlations, NLPCA is

not as stable as a PCA solution (Linting *et al,* 2007a; Linting *et al,* 2007b), where stability is defined as "[…] the degree of sensitivity of an analysis to changes in the data" (Linting *et al,* 2007b). Despite this, NLPCA does not have the same restrictions as polychoric correlations with regard to the number of response categories and missing data, enabling retention of as much information as possible. Both polychoric correlations and NLPCA are limited by the focus in the literature on guidance centered on the performance of factor analysis with continuous data (Gaskin and Happell, 2013). Expert advice was sought through discussion with Dr. Mariëlle Linting, Associate Professor at Leiden University in the Netherlands, who recommended NLPCA as appropriate for the purposes of this study (Linting, 2015, email communication). Although not originally included in the plan for this research, given the uncertainty in the literature it was decided to run the initial analysis twice using both NLPCA and PCA in parallel to test their appropriateness by exploring the differences between the outputs of the two approaches. Following comparison and determination of the most appropriate analysis method, further development of the RA stiffness PROM and preliminary validity testing would be performed (Chapter 9).

8.2 Aims and objectives

The overall aim of this study was to use the responses to the candidate items in a survey of participants to develop and then test the structure of a new RA stiffness PROM. The specific objectives of this chapter were:

- To present the details of the data collection survey
- To compare PCA and NLPCA
- To decide whether PCA or NLPCA is the most appropriate method for analysis of this dataset to take forward into Chapter 9 for further analysis to develop the structure of a new RA stiffness PROM

8.3 Methods

8.3.1 Postal survey

Although computer assisted approaches such as internet surveys are appealing as they can limit measurement error and provide a time efficient survey option (Streiner and Norman, 2008; Groves *et al*, 2009), daily computer use in the target population (mostly >65 years) is only 42% (Office of National Statistics, 2014). Recruiting patients in clinic or using a postal survey enables completion of the questionnaire in paper format, which is likely how it would be used in research or clinical settings. The clinic system in the local rheumatology department is one where RA patients initiate

appointments when required rather than being offered them routinely (Hewlett *et al,* 2005). Therefore when patients attend clinic appointments they are likely to be more unwell and so recruitment would sample patients with a narrower range of disease activity than in a postal survey. A postal survey was therefore used to sample a wider range of patients, and allow more rapid recruitment, as demonstrated in two previous, locally completed survey studies (Nicklin, 2009; Sanderson, 2009).

8.3.2 Patient identification and sampling

Ethics approval was obtained from the University of the West of England REC (HAS/14/10/35) and from Wales REC 4 following proportionate review (14/WA/1162). Inclusion criteria were a confirmed diagnosis of RA (Arnett *et al,* 1988; Aletaha *et al,* 2010) and aged ≥18 years. Exclusion criteria were participation in earlier studies within this research and lack of capacity to consent. The absence of current RA stiffness was not an exclusion criteria, as any PROM developed needs to be able to differentiate those with stiffness from those without.

Recruitment took place between January and March 2015 from NHS patient databases at the BRI and Weston General Hospital. At the BRI, the new patient pathway and direct access databases were used thus providing a range of disease duration and likely disease activity. Databases were checked for duplicate patient entries. Each patient was assigned a random number and patients were then selected in sequence until recruitment was completed. For each patient, medical records were checked to ascertain the inclusion criteria were met and they were not recently deceased. Each patient was assigned a unique identification code.

Questionnaire packs (Section 8.3.3) were sent out in batches so as to meet the recruitment target (Section 8.3.2.1) and to ensure ease of monitoring recruitment. Patients were sent a questionnaire pack via post, inviting them to complete and return the pack in the enclosed prepaid envelope. If questionnaire packs were not returned within three weeks, a reminder pack was sent.

8.3.2.1 Sample size

The sample size for this study was based on considerations relating to the sampling and analysis approaches. A 50% response rate was expected given previous, local research on a similar patient population (Wilson, 2016). Published sample size recommendations for factor analysis are inconsistent (MacCallum *et al,* 1999). Recommendations include the minimum number of participants and the ratio of

participants to items. The recommended minimum number of participants vary. Some describe 100 as acceptable (Gorsuch, 1983) while others suggest that 100 is poor (Comrey and Lee, 1992), 300 is acceptable (Tabachnick and Fidell 2001), or good (Comrey and Lee, 1992), and >1000 is excellent (Comrey and Lee, 1992). Recommendations regarding the ratio of participants to items also differ, ranging from five to 10 (Tinsley and Tinsley, 1987) to 10-15 (Pett, Lackey and Sullivan, 2003) participants per item. With a maximum of 45 items (39 draft and six traditional stiffness items) that could be included in the analysis, approximately 225 completed and returned questionnaire packs would be required to meet minimum recommendations of five participants per item. Therefore, recruitment targeted approximately 450 participants.

8.3.3 Questionnaire pack

A questionnaire pack was developed considering recommendations regarding maximising questionnaire return rates (Streiner and Norman, 2008), and with input from the supervisory team and patient partners. The pack contained a site-specific invitation letter, a patient information sheet, a questionnaire booklet, and a stamped return envelope. The questionnaire pack (Appendix V) included a brief introduction, two consent forms (one of which could be removed from the pack for the participant to keep for their own records) and items in three key sections; 1) clinical items; 2) stiffness items; and 3) demographic items. Table 8.1 provides description and justification of each questionnaire pack component. It also includes the full and abbreviated wording for each stiffness item as included in the final questionnaire pack (Appendix V). In subsequent reference to stiffness items, the full item wording is used in the text while abbreviated item wording is used in tables. The final questionnaire pack was reviewed on two occasions, once with the supervisory team and once with a patient partner (GB).

Two versions of the questionnaire pack were developed with the items in different orders to attempt to combat bias as a result of order effect (Oppenheim, 1992). Questionnaire pack A was printed in blue and questionnaire pack B was printed in green for clarity. Questionnaire pack A was ordered as described above while in questionnaire pack B, half of the clinical items were included before the stiffness items and the rest were included before the demographic items. As the content was identical, only questionnaire pack A is presented (Appendix V). Alternate packs were sent to the randomly sequenced patients (Section 8.3.2).

Table 8.1: Questionnaire pack contents, full item wording, abbreviated item wording and rationale

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
1	Patient Global Assessment (PtG VAS) (van der Heijde *et al,* 1993)	Considering all the ways that your arthritis affects you, mark an X on the scale for how well you are doing	NA	The PtG is a 10cm VAS which asks patients to indicate how well they are doing with their arthritis from 0 (very well) to 10 (very badly). The PtG is taken from the validated DAS28. As has been identified, different wordings of the PtG are often used (Section 1.4). The specific wording used in this study was consistent with the arthritis impact measurement scales (AIMS) (Meenan, Gertman and Mason, 1980).	Although often classified as a measure of disease activity (Anderson *et al,* 2011), it is also suggested that the PtG captures general health or arthritis impact (Kalyoncu *et al,* 2009; French *et al,* 2013). The PtG wording used is consistent with the AIMS (Meenan, Gertman and Mason, 1980) which aims to capture impact. This item is commonly used clinically and the specific PtG wording is used locally (French *et al,* 2013) thus will be recognisable to patients. It will be used for describing the patient sample and exploring the validity of the draft stiffness items.
	Pain NRS (Farrah *et al,* 2001; Hawker *et al,* 2011)	Please circle the number which shows how much pain you have had in the past 7 days	NA	This pain item is an 11-point NRS which asks patients to indicate how much pain they have experienced within the last week from 0 (no pain) to 10 (worst possible pain).	The pain NRS will provide a patient report of an important patient symptom. Previous research and earlier qualitative studies suggested a relationship between pain and stiffness therefore, this item will be used for describing the patient sample and exploring the validity of the draft stiffness items. The timeframe for this item varies although a 24 hour timeframe is most commonly used (Hawker *et al,* 2011). A seven day timeframe has

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
					been used here for consistency with other items in the questionnaire pack.
	Bristol Rheumatoid Arthritis Fatigue Severity NRS (BRAF-NRS) (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b)	Please circle the number which shows your average level of fatigue during the past 7 days	NA	The BRAF-NRS severity item is a patient self-report of fatigue severity over the past seven days. The item asks patients to circle the number representing their average level of fatigue from 0 (no fatigue) to 10 (totally exhausted).	The BRAF-NRS severity item will provide a patient report of fatigue severity. Fatigue is an important patient symptom which has been suggested to relate to stiffness therefore, this item will be used for describing the patient sample and exploring the validity of the draft stiffness items.
	Flare question (Preliminary Flare Questionnaire (PFQ)) (Bykerk *et al,* 2012; Bykerk *et al,* 2014b)	Are you having a flare (flare-up) of rheumatoid arthritis at this time?	NA	This item was taken from the PFQ (Bykerk *et al,* 2012; Bykerk *et al,* 2014b). It asks patients to report whether they consider their RA to be in flare at the present time using a 'Yes' or 'No' response.	The flare question will provide a patient report of whether patients consider their RA to be in flare at present. As qualitative work indicated that stiffness related to flare this item will be used for describing the patient sample and exploring the validity of the draft stiffness items.
	Modified Health Assessment Questionnaire (MHAQ) (Pincus *et al,* 1983)	NA	NA	The MHAQ is short version of the original 20 item HAQ (Fries *et al,* 1980) assessing disability. The MHAQ includes 8 items, one from each of the 8 categories in the HAQ (dressing and grooming, rising, eating, walking, hygiene, reach, grip, activities). However unlike the HAQ the MHAQ does not address the use of aids or	The MHAQ will provide a patient report of perceived disability. It will be used for describing the patient sample and exploring the validity of the draft stiffness items. Although the original HAQ (Fries *et al,* 1980) was used in earlier studies, the MHAQ will be used here because it forms part of the PDAS2 scoring

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
				assistive devices. Patients rate each question with a score between 0 (without any difficulty) and 3 (unable to do) with higher total scores indicating worse function and greater disability. These scores are summed and averaged to give a total MHAQ score between 0-3.	algorithm (see below) and its condensed format reduces participant burden in a large questionnaire pack.
	Patient-based Disease Activity Score (PDAS2) (Choy *et al,* 2008; Choy *et al,* 2015)	NA	NA	The PDAS2 is a composite measure to assess patient reported RA disease activity. The PDAS2 includes four items; a PtG VAS, an EMS duration item with 6 response options (see traditional stiffness duration item D (item no. 6.2)), a 28-SJC on a mannequin displaying individual joints, and the MHAQ (see above). A simplified PDAS2 algorithm can also be calculated without EMS (Choy and Leung, 2016). The PDAS2 without EMS was utilised in this study to avoid circular reasoning.	The PDAS2 will provide a patient report of disease activity. Patient reported assessment is essential given that the study is based on a survey where clinician based assessments is not possible. It will be used for describing the patient sample and exploring the validity of the draft stiffness items. The PDAS2 was developed to provide comparable information to that gained in the DAS28 (Choy *et al,* 2008; Choy *et al,* 2015; Anderson *et al,* 2011) and has demonstrated strong correlations with DAS28 (r_s=0.76, *p* not reported) (Choy *et al,* 2008).
2	Draft stiffness items	Please circle the number that best describes the severity of your RA stiffness over a usual week when you are not in a flare?	Draft item response shift	The draft stiffness items have been developed during previous qualitative studies, as described in earlier chapters.	The 45 draft stiffness items are included to develop the structure of the final combination of items that work best together, followed by preliminary validity testing (Chapter 9).

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
		Do you have any joints that are permanently stuck?	Draft item stuck joints		
		Over the past 7 days when have you experienced RA stiffness?	Draft item timing		
		Have you experienced RA stiffness in your joints over the past 7 days?	Draft item in joints		
		Over the past 7 days have you experienced RA stiffness all over?	Draft item all over		
		Over the past 7 days has your RA stiffness been different to usual for you?	Draft item different to usual		
		Over the past 7 days has your RA stiffness been as variable (coming and going) as usual for you?	Draft item variable		
		Over the past 7 days have you experienced RA stiffness after a period of immobility (for example, after sitting for a while)?	Draft item after immobility		
		Have you experienced RA stiffness in your body	Draft item in body		

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
		(outside of your joints) over the past 7 days?			
		Has RA stiffness affected your sleep?	Draft item sleep		
		Has RA stiffness made it difficult to dress or undress yourself?	Draft item dress		
		Has RA stiffness made it difficult to wash yourself (for example, have a shower)?	Draft item wash		
		Has RA stiffness made it difficult to carry out your responsibilities or commitments?	Draft item responsibilities		
		Has RA stiffness made it difficult to do your daily tasks or activities?	Draft item daily tasks		
		Has RA stiffness made it difficult to chew?	Draft item chew		
		Has RA stiffness made it difficult to do hobbies or activities you enjoy?	Draft item hobbies		
		Has RA stiffness made it difficult to get out of bed?	Draft item get out of bed		

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
		Has RA stiffness made it difficult to get up after sitting for a while?	Draft item get up after sitting		
		Have your daily tasks and activities required more effort because of RA stiffness?	Draft item effort		
		Has RA stiffness made you slower (for example, unable to do things quickly)?	Draft item slower		
		Has RA stiffness made it difficult to do fine movements (for example, write with a pen)?	Draft item fine movement		
		Has RA stiffness made it difficult to grip or hold things?	Draft item grip		
		Has RA stiffness made it difficult to open and close your fist?	Draft item open/close fist		
		Has RA stiffness reduced your strength to do tasks?	Draft item strength		
		Has your movement been restricted because of RA stiffness?	Draft item movement		

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
		Has RA stiffness made it difficult to balance without physically supporting yourself?	Draft item balance		
		Have you had to concentrate to move your body because of RA stiffness?	Draft item concentrate		
		Have you felt frustrated because of RA stiffness?	Draft item frustrated		
		Have you felt worried or concerned because of RA stiffness?	Draft item worried		
		Have you felt self-conscious because of RA stiffness?	Draft item self-conscious		
		Has it taken you longer to do your daily tasks or activities because of RA stiffness?	Draft item take longer		
		Have you had to change your plans or behaviour because of RA stiffness?	Draft item change plans		
		Have you had to work around your RA stiffness (or do things in a different way)?	Draft item work around		

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
		Have you needed help (from others or gadgets) because of RA stiffness?	Draft item need help		
		Please circle the number that best describes the impact that RA stiffness has had on your life over the past 7 days	Draft item impact		
		Please circle the number that best describes the severity of your RA stiffness over the past 7 days	Draft item severity		
		Please circle the number that best describes how important RA stiffness has been in your life over the past 7 days	Draft item importance		
		Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days	Draft item coped		
		How much of the stiffness you have reported in the questions above is about joints that are permanently stuck?	Draft item stuck joints B		

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
	Traditional stiffness items	How would you describe the overall level of morning stiffness you have had from the time you wake up?	Traditional item severity A	This is an 11-point NRS which asks patients to indicate the overall level of MS from waking from 0 (no stiffness) to 10 (very severe stiffness). The specific wording comes from the Bath Ankylosing Spondylitis Disease Activity Index (BASDI) which is a validated composite score used in AS (Garrett *et al,* 1994). The NRS anchors have been used in previous studies (e.g. Rhind, Unsworth and Haslock, 1987; Hazes, Hayton, and Silman, 1993).	As described earlier (Section 6.3.2.1.4) it has been suggested that stiffness assessed in the form of severity has better measurement properties than duration (e.g. Westhoff *et al,* 2008; Lie *et al,* 2014) despite being less commonly assessed (Cutolo, 2011). Poor definition of stiffness items is common in the literature (Section 2.3.1). Wordings, response options and formats of traditional severity items used in this study have been based on defined items used in the literature where possible. Adjustments have been in circumstances where items were not defined. For example the use of the wording from the BASDAI (Garrett *et al,* 1994) which has also been used in studies involving RA patients (e.g. Lie *et al,* 2014). The stiffness severity item B which was based on a VAS has not been included here for reasons discussed in Chapter 7. These traditional stiffness severity items will be included in further testing to identify the smallest combination of items that perform most effectively to assess stiffness.
		Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week	Traditional item severity G	This is an 11-point NRS which asks patients to indicate stiffness during the last week from 0 (no stiffness) to 10 (extreme stiffness). This item is taken from the PFQ and has been suggested to be able to distinguish patients reporting being in flare from those reporting not being in flare (Bykerk *et al,* 2012; Bykerk *et al,* 2014b). This item was not originally included for testing (Chapter 6) however, as detailed previously (Section 7.5.2.4) the wording for this item is clearly defined and fits with patient descriptions of stiffness therefore it has been included subsequently.	

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
		How would you describe the overall level of morning stiffness you have had from the time you wake up?	Traditional item severity C	This item asks patients to indicate the overall level of MS from waking using the same wording as item 12.1, derived from the BASDAI (Garrett *et al,* 1994). The response options are a 5-point Likert scale from no stiffness to very severe stiffness, as used in previous studies (e.g. Rhind, Unsworth and Haslock, 1987).	
		Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?	Traditional item duration D	This item asks patients to indicate whether their joints were stiff on waking today, and if so the subsequent duration within six possible response options between 'less than 30 minutes' to 'all day'. This item has been included in a number of validated composite scales including the RADAI (Stucki *et al,* 1995; Fransen *et al,* 2000; which was modified from RADAR (Mason *et al,* 1992)), and the PDAS2 (Choy *et al,* 2008; Choy *et al,* 2015).	Stiffness assessed in the form of duration is recommended in the ACR guidelines for RA management (American College of Rheumatology subcommittee on rheumatoid arthritis guidelines, 2002) and is the most common stiffness assessment method in trials (Cutolo, 2011). This is despite the suggestion that stiffness assessed by severity has better measurement properties (e.g. Westhoff *et al,* 2008).
		How long does your morning stiffness last from waking until maximum improvement occurs?	Traditional item duration E	This item asks patients to indicate the duration of MS from waking to maximum improvement. The wording is in accordance with ARA guidelines (Arnett *et al,* 1988), has been employed in other research (e.g. Vliet Vlieland *et al,* 1997; Khan *et al,* 2009), and was found to be most effective compared to other similar items	Traditional duration items have used a number of different wordings and formats, although the exact wording and format are rarely described in studies (Section 2.3.1). Wordings, response options and formats of traditional severity items used in this study have been based on defined items used in the literature where possible.

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
				(Hazes *et al,* 1994). The item provides three possible response options ('up to 1 hour', '1-3 hours', and 'more than 3 hours'). These options had been used in categorisation of responses in the literature (Hazes *et al,* 1994).	As with the traditional stiffness severity items, these traditional duration items will be included in further testing to identify the smallest combination of items that perform most effectively to assess stiffness.
		How long does your morning stiffness last from waking until maximum improvement occurs?	Traditional item duration F	This item asks patients to indicate the duration of MS from waking to maximum improvement and is consistent with the wording of item 12.2. The item provides response options in minutes and/or hours. This format has been used in previous studies (e.g. Rhind, Unsworth and Haslock, 1987; Vliet Vlieland *et al,* 1997).	
3	Gender	Are you male or female?	NA	Simplified wording for this item was adapted from the 2011 census	The demographic items are required to enable description of the sample. Specifically, it is important to be able to compare the sample in this study to the samples involved in questionnaire development and also compare responders and non-responders to inform the generalisability of the study findings.
	Age	What is your date of birth?	NA	Simplified wording for this item was adapted from the 2011 census	
	Disease duration	Approximately how long have you had rheumatoid arthritis (RA)?	NA	Wording of this item was adapted from a recent survey conducted within the department	
	Current RA medications	What medications are you taking for your RA?	NA	Wording of this item was adapted from a recent survey conducted within the department	

Section	Item/scale	Full wording	Abbreviated wording	Source/content	Rationale
	Co-morbidities	Do you have any other medical conditions for which you are receiving treatment?	NA	Wording of this item was adapted from a recent survey conducted within the department	
	Work status	What is your work status?	NA	Simplified wording for this item was adapted from the 2011 census and a recent survey conducted within the department	
	Education level	What is your level of education?	NA	Simplified wording for this item was adapted from the 2011 census and a recent survey conducted within the department	
	Postcode	What is your postcode?	NA	Wording of this item was adapted from a recent survey conducted within the department	

8.3.3.1 Index of multiple deprivation

Postcode data were collected to obtain a measure of deprivation. Socioeconomic factors were recognised as an important topic in relation to outcome measures in the OMERACT equity special interest group (O'Neil *et al,* 2014). These data, alongside other demographic data, characterised the patient sample in which the new stiffness PROM was developed. This information will be useful to identify areas for future development and validation of any new PROM.

Deprivation information was collected using the index of multiple deprivation (IMD). The IMD is a measure of deprivation calculated from seven weighted domains; income (22.5%), employment (22.5%); health and disability (13.5%); education, skills and training (13.5%); barriers to housing and services (9.3%); crime (9.3%); and living environment (9.3%) (Department for Communities and Local Government, 2011). The IMD represents the level of deprivation in specific geographical areas or 'lower layer super output areas' (LSOA's). There are 32,482 LSOA's in England which are ranked from one (most deprived) to 32,482 (least deprived). IMD scores represent the whole LSOA and may not be directly applicable to an individual within that area. GeoConvert (GeoConvert, 2007), a free online tool that can match postcodes to measures of deprivation was used to generate IMD 2010 scores. IMD scores for the whole sample were converted into categories with category one representing the least deprived scores (lowest 20%) and category five the most deprived scores (81-100%), (other categories; category two (21-40%); category three (41-60%); category four (61-80%)). These categories are therefore relative, limited to the geographical area in which the participants live.

8.4 Analysis and results

A 4-stage analysis plan was devised to ensure an organised and directed approach to data analysis. Stage 1 involved data cleaning and descriptive statistics. Stage 2 involved identification of an appropriate analysis method. Stage 3 and 4 focused on the development of the new RA stiffness PROM and preliminary validity testing and are presented in Chapter 9. The description of each analysis stage is followed directly by its results.

8.4.1 Stage 1 analysis: Data cleaning and descriptive statistics

An Excel spreadsheet was developed for study management to monitor the sending and return of questionnaire packs. SPSS for Windows (version 21) was used for data

management and analysis. Each questionnaire pack was reviewed by hand by the researcher. Data gathered using existing instruments (PDAS2 and MHAQ) were scored according to the authors' instructions using Excel formulae. All data were then entered into SPSS according to a codebook which was developed to aid consistent data input (Pallant, 2010). Within SPSS, descriptives and frequencies, correlations, and the factor analysis and categorical principal component analysis (CATPCA) programmes were used.

8.4.1.1 Data cleaning

The dataset was inspected for error as a result of data input. Frequency distributions were generated for each variable and inspected for minimum and maximum values, and missing values. Identified potential errors were checked in the original questionnaire pack and if necessary corrected in the database. Following correction, frequencies and descriptives were repeated for each variable.

8.4.1.2 Descriptive statistics

Descriptive statistics were performed to enable characterisation of the sample and to compare responders and non-responders. Given the nature of these data, medians and interquartile ranges (IQR) were reported and variables were described using frequencies and percentages. As some variables were continuous in nature (e.g. age), or could be classed as interval (e.g. pain assessed on an 11-point NRS), the extent to which these variables met the assumptions of normality was explored using measures of skew and kurtosis, and the Kolmogorov-Smirnov statistic. However, in large samples, these tests are often too sensitive and inspection of the histogram is recommended (Tabachnick and Fidell, 2007). Where variables demonstrated reasonably normal distributions and values on the Q-Q plot fell near to or on the straight line, mean and standard deviations were also reported.

8.4.1.3 Missing data

Errors such as missing information are common in self-completed surveys (Silman and Macfarlane, 2002), although the pattern of missing data may be more relevant than the amount (Tabachnick and Fidell, 2007). For this study, it was important to identify any items with large amounts of missing data as this would be taken into account during decision making about the usefulness of items during analysis stages 2 and 3. It was also relevant for analyses where decisions are required regarding the use of listwise (include only cases with a complete dataset across all variables) or

pairwise (include only cases with a complete dataset for specific variables) treatment of missing values (Pallant, 2010). Frequency outputs and pattern analysis were used to identify missing values. Missing values up to 5% were defined as acceptable as this amount is not likely to cause serious problems in large datasets (Tabachnick and Fidell, 2007). Items that were marked by participants as not being relevant were coded differently (8888) from items that were left blank (9999). As the reason for an item being left blank is unknown these were always treated as missing (9999).

8.4.2 Stage 1 results: Data cleaning and descriptive statistics

8.4.2.1 Data cleaning

Seven variables had incorrectly entered values which were corrected. Histograms and investigation into normality (e.g. Q-Q plots) for non-stiffness variables are available in Appendix W. Frequency and distribution graphs for all draft stiffness items are presented in Section 8.4.4.1.2.

8.4.2.2 Descriptive statistics

8.4.2.2.1 Study population

Of 645 questionnaire packs sent in seven batches, 197 were returned without a reminder, a further 80 were returned following a reminder and two were duplicates, giving an overall response rate of 43.1% (n=277) (Figure 8.1).

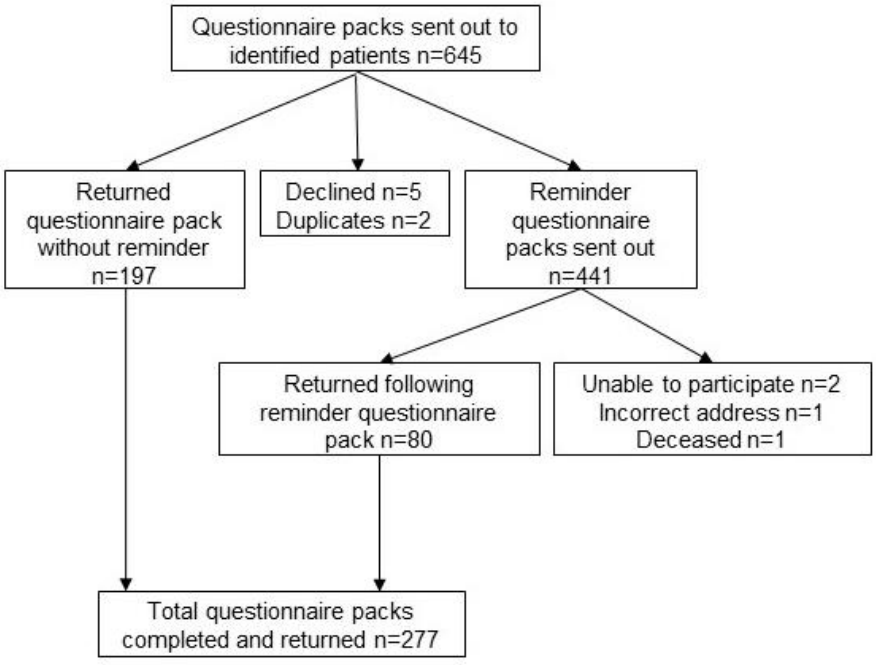


Figure 8.1: Recruitment flow diagram

8.4.2.2.2 Responder and non-responder demographic information

Frequencies and percentages for gender, age and social deprivation were similar between responders (participants) and non-responders (Table 8.2).

Table 8.2: Demographic information across whole sample

		Whole sample (n = 643)*	Responders (n = 277)*	Non-responders (n = 366)*
Gender count (%)	Female	465 (72.3%)	186 (67.1%)	279 (76.2%)
	Male	178 (27.7%)	91 (32.9%)	87 (23.8%)
Age	Mean (SD)	62.0 (14.0)	63.9 (12.4)	60.5 (15.0)
	Median (IQR)	64.0 (53.0-73.0)	65.0 (55.3-74.0)	62.0 (51.0-72.0)
Social deprivation (IMD 2010 score) count (%)	1 (Least deprived)	273 (43.1%)	134 (49.1%)	139 (38.5%)
	2	197 (31.1%)	84 (30.8%)	113 (31.3%)
	3	80 (12.6%)	28 (10.3%)	52 (14.4%)
	4	52 (8.2%)	15 (5.5%)	37 (10.2%)
	5 (Most deprived)	32 (5.0%)	12 (4.4%)	20 (5.5%)

*GeoConvert unable to match 9 postcodes (whole sample n=634; responders n=273; non-responders n=361)

8.4.2.2.3 Participant disease and medication demographic information

Participants were representative of an RA population; mostly female and >45 years of age. Gender and age were similar to reports from other recent surveys (74% female, mean age=63.5 years, Wilson *et al,* 2015; 75% female, mean age=60 years, Hammond *et al,* 2015). Participants had a range of disease duration (n=271, median=6 years, IQR=3.0-15.0 years) and most frequently reported taking DMARDS, while 20% were taking biologics. The majority (80.2%) reported taking medication for other comorbidities (Table 8.3). A small number of participants reported that they had conditions where stiffness was a feature, including OA (n=4), PMR (n=1), AS (n=1). This will be addressed in the discussion.

Disease activity was assessed using PDAS2 (Choy *et al,* 2008, Choy *et al,* 2015). PDAS2 scores demonstrated a range of disease activity (range 2.7-6.9, median=4.1, IQR=3.2-5.0). Categorisation of scores was based on established cut-offs from previous validation work (Leung *et al,* 2012). Although 38.5% (n=106) were in remission, over a quarter of participants had high disease activity (n=80, 28.9%).

Table 8.3: Participant demographic information

Variable	Frequency	Percent	Missing
Gender			0
Male	91	32.9%	
Female	186	67.1%	
Age			0
<44	16	5.8%	
45-64	119	43.0%	
65-74	78	28.2%	
>74	64	23.1%	
Disease duration			6
≤2	45	16.6%	
3-<5	60	22.1%	
5-<10	58	21.4%	
10-20	68	25.1%	
>20	40	14.8%	
Medications			8
Analgesics and NSAIDS	104	38.7%	
DMARDs	219	81.4%	
Glucocorticoids	98	36.4%	
Biologics	55	20.4%	
Comorbidities			33
None	47	19.3%	
1 other condition	98	40.2%	
2 other conditions	64	26.2%	
3 or more other conditions	35	14.3%	
Disease activity (PDAS2)			0
Remission (<3.8)	106	38.3%	
Low disease activity (3.8-4.5)	65	23.5%	
Moderate disease activity (4.6-5.0)	26	9.4%	
High disease activity (>5.0)	80	28.9%	
Disability (MHAQ)			0
Normal (<0.3)	112	40.4%	
Mild functional loss (0.3-<1.3)	120	43.3%	
Moderate functional loss (1.3-1.8)	32	11.6%	
Severe functional loss (>1.8)	13	4.7%	
Patient global assessment (PtG)			2
0-39 (mild)	141	51.3%	
40-69 (moderate)	89	32.4%	
70-100 (severe)	45	16.4%	
Pain			0
0-3 (mild)	104	37.5%	
4-6 (moderate)	87	31.4%	
7-10 (severe)	86	31.0%	
Fatigue (BRAF-NRS)			0
0-3 (mild)	75	27.1%	
4-6 (moderate)	94	33.9%	
7-10 (severe)	108	39.0%	

Patient reports of perceived disability were captured using the MHAQ (Pincus *et al,* 1983), demonstrating a range of disability (range 0-2.5, median=0.5, IQR=0.1-1.1). Categorisation of scores was based on established cut-offs from previous validation

work (Maska, Anderson and Michaud, 2011). Mild functional loss (n=120, 43.3%) was most frequently reported. A range of scores in relation to pain (NRS, median=5.0, IQR=2.0-7.0), fatigue (NRS, median=6.0, IQR=3.0-7.0) and PtG (100cm VAS, median=36.0, IQR=16.0-59.0) were also reported, and 102 (37%) participants reported being in flare. These variables will be explored further in Chapter 9.

8.4.2.2.4 Participant sociodemographic information

Most participants were retired (n=146, 52.9%) but 80 (29.0%) were in paid work (Table 8.4). A small group of participants reported a combination of responses (n=17, 6.2%), most commonly retired and receiving incapacity benefits (n=11). Of the 257 participants who reported their education level, 54.1% had a school education. A range of IMD scores were included in the sample (range 0.99-69.65, median=14.2, IQR=9.4-25.7), although most participants resided in areas of low deprivation (Table 8.2).

Table 8.4: Participant sociodemographic information

	Frequency	Percent	Missing
Work status			1
Student	0	0.0%	
Paid work	80	29.0%	
Homemaker	4	1.4%	
Unemployed	7	2.5%	
Retired	146	52.9%	
Receiving incapacity benefits	22	8.0%	
Combination of responses	17	6.2%	
Education level			20
Did not complete school	2	0.8%	
School education	139	54.1%	
College/apprenticeship	62	24.1%	
University	54	21.0%	

8.4.2.3 Missing data

Overall the number of missing values for the draft stiffness items was small with only 232 of the 12,465 total values missing (1.9%), although 40 of the 45 draft stiffness items (88.9 %), and 104 of the 277 participants (37.6%) had at least one missing value. Pattern analysis identified only two items with >5% missing data. Both were traditional stiffness items asking 'How long does your morning stiffness last from waking until maximum improvement occurs?', one on a ordinal scale (n=39, 14.1%), the other with minutes and hours as the response option (n=56, 20.2%). Most participants had no missing stiffness values (n=173, 62.5%). Of the 104 participants with missing values, most (n=80, 76.9%) had <5% (representing one or two missing

values). The remainder had three to eight missing values (n=24, 23.1%), except one participant who had 26, accounting for 11.2% of all missing stiffness values.

For non-stiffness items, pattern analysis identified only one item (13.7, education level) with missing data over >5% (n=20, 7.2%). Missing data was therefore generally limited to specific items or participants. As information on missing data was being considered in item suitability decision making (Section 8.4.4.1.1), data were not imputed. Given the spread of missing data across approximately one third of participants, pairwise treatment of missing data was used to maximise the dataset.

8.4.3 Stage 2 analysis: Identification of an appropriate analysis method

Given the considerations regarding the appropriateness of PCA for this dataset (Section 8.1.3), it was important to test the two proposed approaches and select the most appropriate method. To do this, both PCA and NLPCA were performed in parallel. This involved five steps: 1) assess item suitability; 2) preliminary analysis; 3) component extraction; 4) component rotation; 5) compare the approaches. The sections that follow describe each step. Aspects relevant to both PCA and NLPCA have been described together, aspects specific to each analytical approach have been addressed separately where relevant.

8.4.3.1 Assess item suitability

The aim of this step was to ensure that items were suitable to take forward for PCA and NLPCA and before moving on to preliminary analysis, a decision was made about whether to retain or remove items. Decisions were based on the acceptability of items for both PCA and NLPCA but removal of items was conservative to ensure that as many items as possible were retained for further analysis. A consideration relevant for both analyses was missing data therefore items with large amounts of missing data were considered for removal.

8.4.3.1.1 PCA

The premise of PCA is the identification of different groups of related items (Pallant, 2010). Items assessing the same concept will correlate with each other whilst items addressing different concepts will correlate poorly. A Pearson's correlation coefficient matrix is one of the first parts of a PCA output and can be used to explore the relationship between items. It is recommended that the correlation matrix is screened for very low (r<0.30) or very high (r≥0.80) correlations (Pett, 1997; Pett, Lackey and

Sullivan, 2003; Field, 2009). However, as multicollinearity (very high correlation) does not affect PCA (Field, 2009), only items with many correlations below <0.3 were considered for removal (Field, 2009). Recognising that this is an arbitrary cut-off (Field, 2009), for the purpose of this study any item which had correlations <0.3 with more than 36 of the 45 items (80%) was considered for removal.

Other considerations are the frequency and distribution of data. Factor analysis requires roughly normal distributions (Field, 2009) and discussion with a statistician (Rosemary Greenwood, Research Design Service) highlighted that items with a large number of responses in one response category only may cause difficulties with analysis and are unlikely to be useful. Frequency tables and distribution graphs were performed for each item and inspected for the spread of responses across categories. Any item with a response category containing 50% or more of the responses to that item was considered for removal.

8.4.3.1.1 NLPCA

Although both NLPCA and PCA generate a correlation matrix, in NLPCA calculations are not based on the correlation matrix as in PCA but are generated directly from the data themselves (Linting and van der Kooij, 2012) so the correlation matrix is not as important. To assess item suitability in NLPCA, the key consideration is the frequency of responses in each category as low frequencies can lead to unstable solutions or inflated influence on the quantification process. The number of observations per category to ensure stability may vary (Linting and van der Kooij, 2012) but a minimum of eight has been recommended (Markus, 1994, in Linting and van der Kooij, 2012). Frequency tables and distribution graphs were therefore used to identify items with response categories with low frequencies.

Overall, the suitability criteria for items across PCA and NLPCA included missing responses, distributions and frequencies, and correlations. These data are consolidated in Section 8.4.4.1.4.

8.4.3.2 Preliminary analysis

This step reviewed the initial output from each analysis method. This provided an indication as to whether the data were appropriate from a statistical perspective, in other words to explore whether the data were factorable (Pett, Lackey and Sullivan, 2003).

8.4.3.2.1 PCA

Two criteria were reviewed: Bartlett's test of sphericity and Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (Kaiser, 1970, in Field, 2009). Bartlett's test of sphericity assesses whether the correlation matrix is an identity matrix, meaning there is no correlation among items, and should be large and significant (Pett, Lackey and Sullivan, 2003; Field, 2009). The KMO is an indicator of the strength of the overall relationship between items (Pett, Lackey and Sullivan, 2003) and is useful given the variable sample size recommendations for factor analysis (Section 8.3.2.1). It is suggested that KMO values between 0.5-0.7 are 'mediocre', values between 0.7-0.8 are 'good', values between 0.8-0.9 are 'great', and values >0.9 are 'superb' (Hutchenson and Sofroniou, 1999). As recommended, the KMO was also inspected for individual items using the anti-image matrices output and items with diagonal elements <0.5 were considered for removal (Field, 2009).

8.4.3.2.2 NLPCA

One of the advantages of NLPCA is the ability to take into account data at different levels of measurement (e.g. ordinal). The specification of the level of analysis defines the freedom allowed in the category quantifications (Linting and van der Kooij, 2012). Whilst the level of analysis does not need to match the level of measurement of a variable, it is a relevant consideration because increased freedom reduces the stability of solutions (Linting and van der Kooij, 2012). The nominal analysis level allows the most freedom and is useful if exploring nonmonotonic relationships (where variables do not increase or decrease at the same rate as each other). The ordinal analysis level allows slightly less freedom, allowing maintenance of category order, but assumes relationships between variables may not be linear. Numeric analysis level allows the least freedom and as in PCA assumes linear relationships between variables (Linting and van der Kooij, 2012).

Expert advice recommended that variables should only be given enough freedom to adequately describe the data, thus initially the lowest level of analysis (nominal) should be specified and gradually restricted if appropriate (Linting, 2015, email communication). Therefore as recommended, the analysis level was initially specified as nominal for all variables but was also tested with more restricted analysis levels. Variance accounted for (VAF) is an important indicator of fit that is represented by eigenvalues (Field, 2009; Linting and van der Kooij, 2012). Eigenvalues are indicators of the size of a component (Field, 2009) as they represent the VAF by all of the items that make up that component (Pett, Lackey and Sullivan, 2003). If VAF does not show

significant improvement as a result of allowing more freedom then a more restricted analysis level would provide more stability, simpler interpretation of the relationship between variables, and less risk of capitalising on chance (Linting and van der Kooij, 2012).

8.4.3.3 Component extraction

This step involved determination of whether separate factors or components were present (the term 'component' is used for consistency across PCA and NLPCA). Components are subsets or groups of the original items and are based on eigenvalues (Pett, Lackey and Sullivan, 2003). Both PCA and NLPCA define the number of components present in the dataset but use slightly different approaches. In NLPCA, the number of components is specified for the analysis but in PCA it is generated by the analysis (Field, 2009; Linting and van der Kooij, 2012). However, as the number of components to retain was not directly relevant to decisions regarding the most appropriate analytical approach to use, this has not been explored in detail here but is discussed in Chapter 9.

8.4.3.4 Component rotation

Unrotated solutions are often not meaningful or easy to interpret (Pett, Lackey and Sullivan, 2003) and rotation produces a simpler and more interpretable solution by maximising high item loadings and minimising low item loadings (Field, 2009). The goal of rotation is a simple structure where each item has a high or meaningful loading on one component, and each component has high or meaningful loadings for only some of the items (Pedhazur and Schmelkin, 1991). Weak loadings have been defined as ≤0.30 (Hair *et al,* 1995) or <0.30 (Comrey and Lee, 1992), but Pett *et al.* (2003) suggest suppressing loadings <0.40 when evaluating outputs. In this study, loadings ≥0.40 are defined as high or meaningful.

There are two types of component rotation; orthogonal and oblique. Within each type of rotation there are different methods, which have different ways of rotating the components (Field, 2009). As the appropriateness of different approaches to rotation have been discussed in the literature (e.g. Pett, Lackey and Sullivan, 2003), both orthogonal (varimax) and oblique (promax) approaches to rotation were tested and compared (Pett, Lackey and Sullivan, 2003; Field, 2009). Varimax is the most common orthogonal rotation method (Pett, Lackey and Sullivan, 2003) which works by maximising high loadings and minimising low loadings to improve interpretability

(Tabachnick and Fidell, 2001; Field, 2009). Promax aims to produce a clear structure but allows correlations between components (Tabachnick and Fidell, 2001).

The options for rotation can be selected during the PCA procedure but, for NLPCA the CATPCA programme does not allow rotation. As recommended (Linting and van der Kooij, 2012), the transformed variables produced in NLPCA were saved and subject to rotation within the factor analysis programme.

8.4.3.5 Comparing the two statistical approaches

This step compared the findings from PCA and NLPCA to inform the decision about the analysis method to take forward to stage 3 analysis. Four aspects of PCA and NLPCA are directly comparable; eigenvalues; component loadings; communalities; and component scores (Linting and van der Kooij, 2012).

As explained earlier, eigenvalues are indicators of the importance of a component (Field, 2009) as they represent the VAF by all of the items that make up that component (Pett, Lackey and Sullivan, 2003). As VAF is an important indicator of fit (Linting and van der Kooij, 2012), eigenvalues and VAF were compared across analyses. Following discussion with the supervisory team it was also considered important that VAF across rotated components was similar to reflect the conceptual importance of the components and enhance stability and robustness of components in further testing. Rotated component loadings demonstrate the correlations between the quantified variables and the components (Linting and van der Kooij, 2012). Given that rotated solutions are more meaningful and easier to interpret (Pett, Lackey and Sullivan, 2003), the rotated solutions were compared across analyses. Communality is the proportion of common variance in a (quantified) variable that is shared with other variables (Field, 2009; Linting and van der Kooij, 2012). As communalities are indicators of shared common variance (Field, 2009), communalities for each item were compared across analyses. Although component scores are consistent outputs across methods, they are particularly useful for interpretive purposes and primarily used for investigation at an individual participant level for example, whether groups of participants with certain characteristics score highly on certain components. Given the purpose of this part of the analysis, this was not felt to be relevant at this stage therefore, component scores were not considered. Therefore, four criteria were used to compare PCA and NLPCA; 1) eigenvalues and VAF; 2) VAF across components 3) component loadings; and 4) communalities.

8.4.4 Stage 2 results: Identification of an appropriate analysis method

8.4.4.1 Assess item suitability

8.4.4.1.1 Missing responses

Due to missing data, both traditional stiffness items asking 'How long does your morning stiffness last from waking until maximum improvement occurs?' using different response options (ordinal scale and minutes and hours) were considered for removal.

8.4.4.1.2 Distribution and frequencies

Frequency and distribution graphs were generated for each of the 45 items, including the draft stiffness items (Figures 8.3-8.41) and traditional stiffness items (Figures 8.42-8.47). Response categories containing over 50% of all responses, or less than eight responses per category are highlighted in red and are listed in Table 8.5.

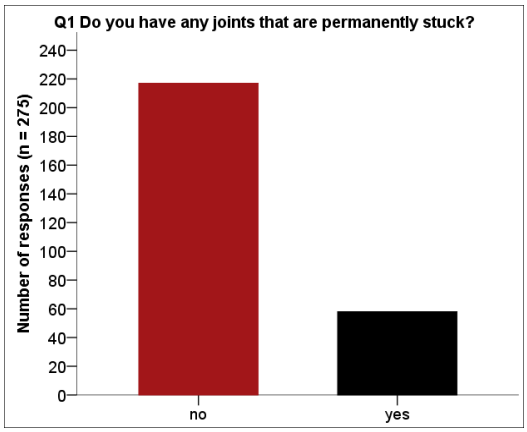


Figure 8.3: Frequency and distribution graph

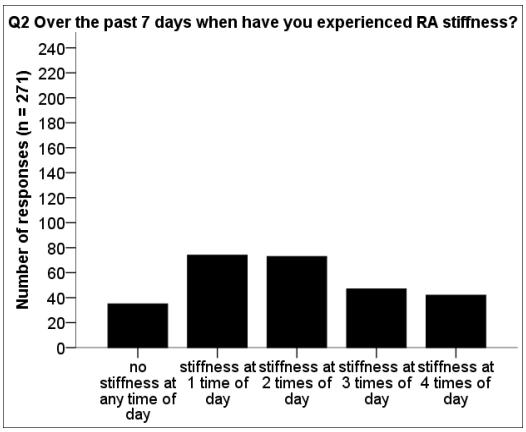


Figure 8.4: Frequency and distribution graph

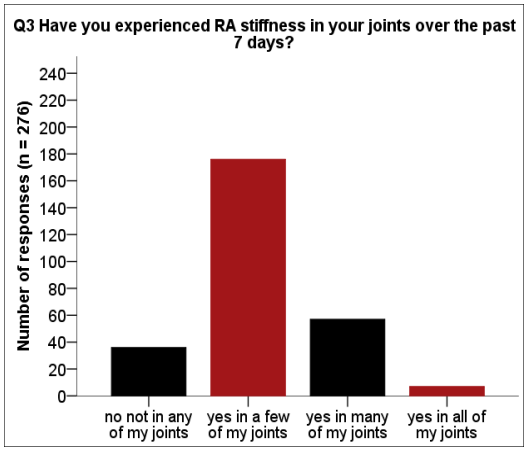


Figure 8.5: Frequency and distribution graph

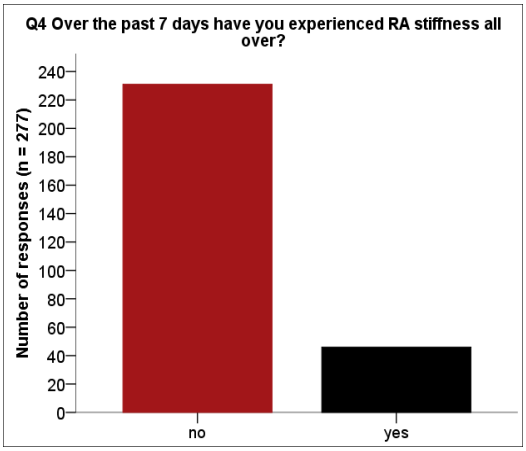


Figure 8.6: Frequency and distribution graph

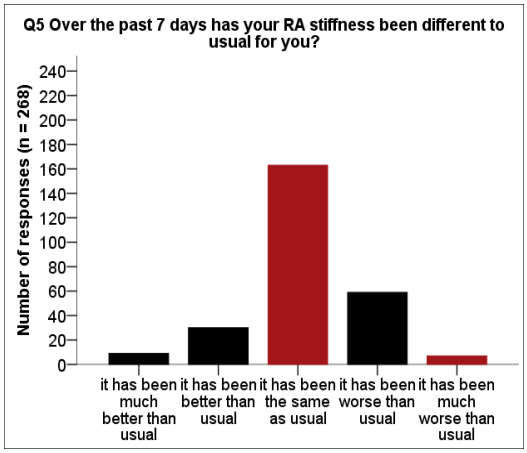


Figure 8.7: Frequency and distribution graph

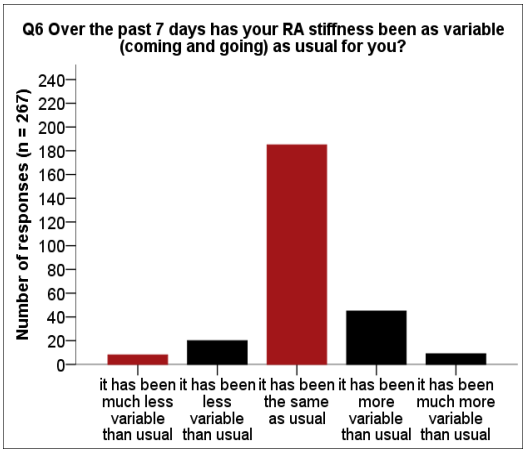


Figure 8.8: Frequency and distribution graph

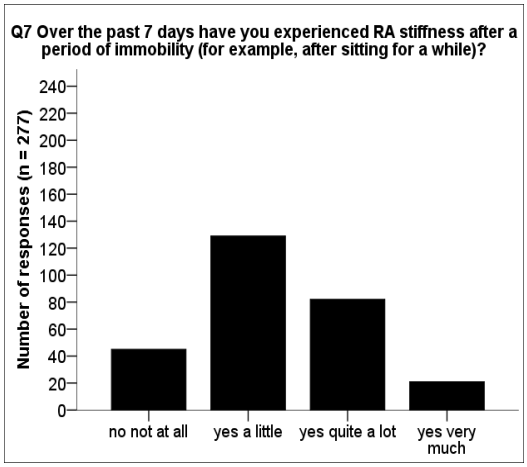


Figure 8.9: Frequency and distribution graph

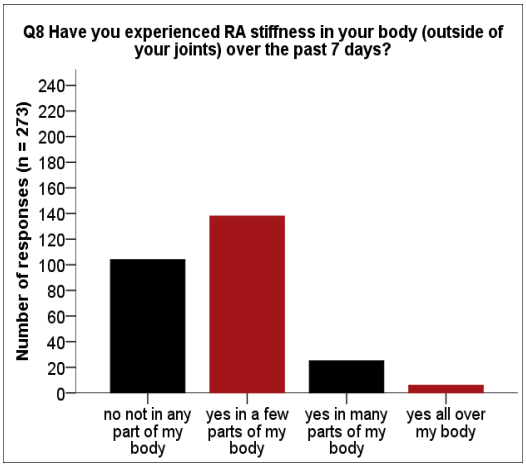


Figure 8.10: Frequency and distribution graph

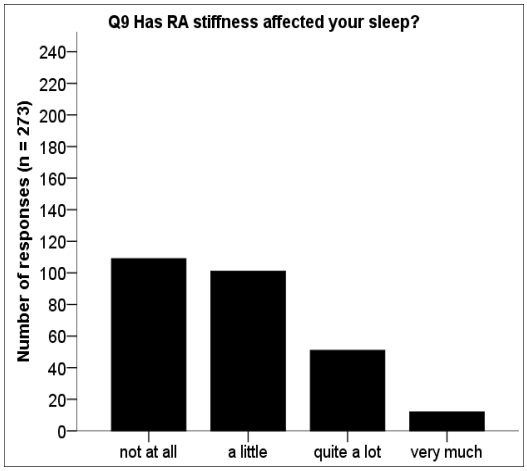


Figure 8.11: Frequency and distribution graph

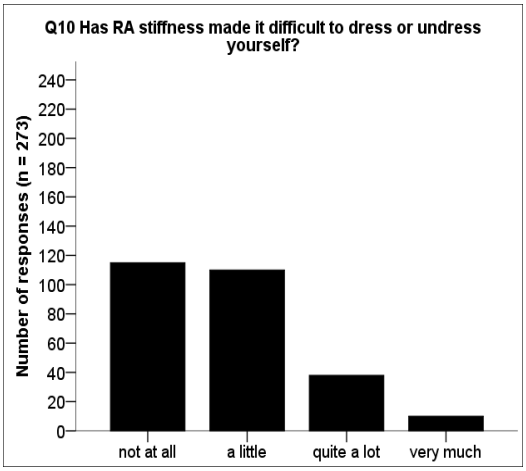


Figure 8.12: Frequency and distribution graph

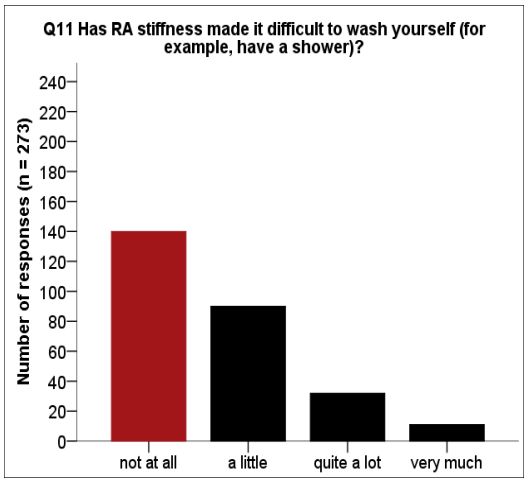


Figure 8.13: Frequency and distribution graph

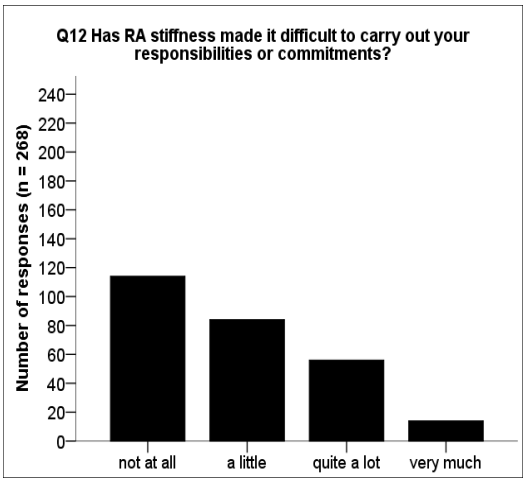


Figure 8.14: Frequency and distribution graph

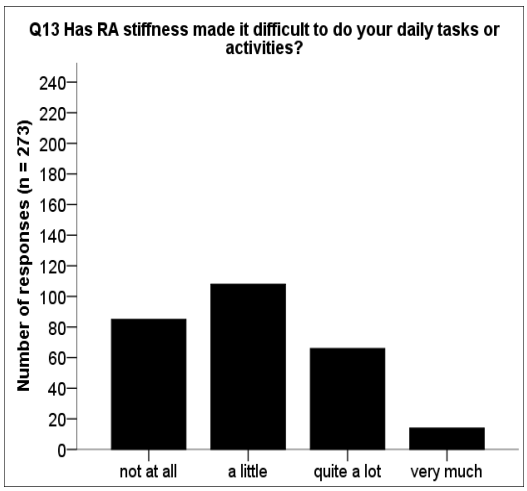


Figure 8.15: Frequency and distribution graph

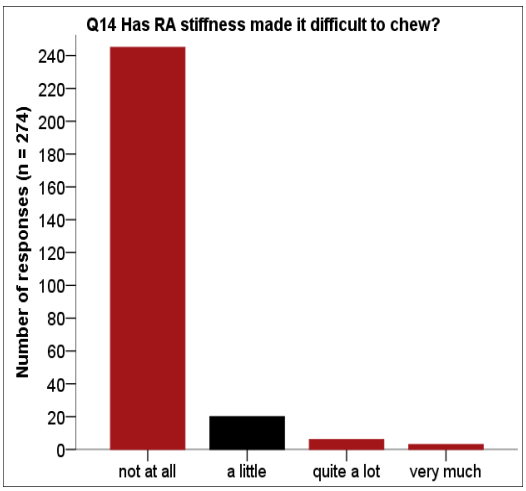


Figure 8.16: Frequency and distribution graph

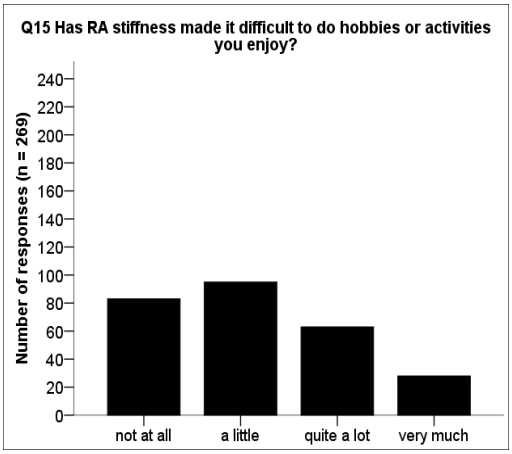


Figure 8.17: Frequency and distribution graph

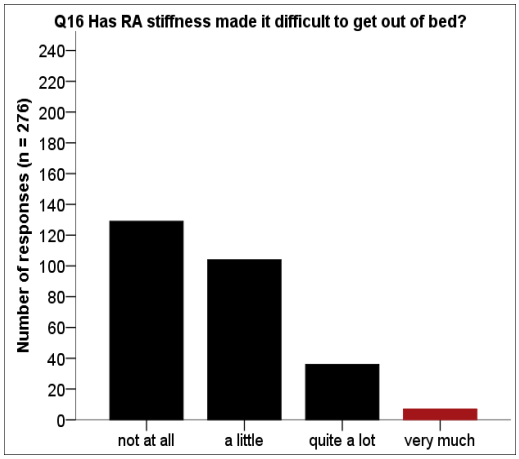


Figure 8.18: Frequency and distribution graph

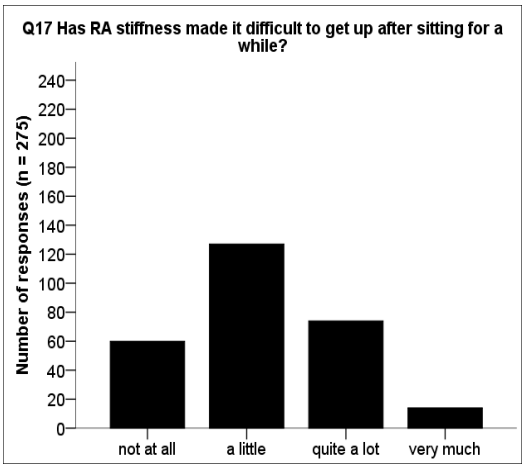


Figure 8.19: Frequency and distribution graph

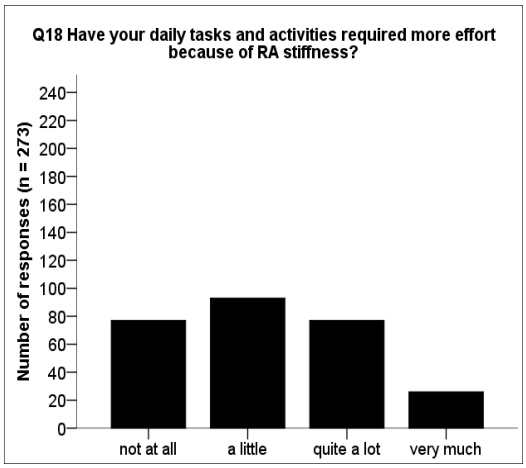


Figure 8.20: Frequency and distribution graph

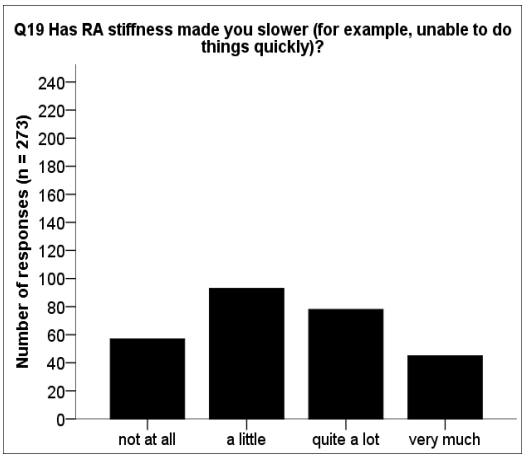


Figure 8.21: Frequency and distribution graph

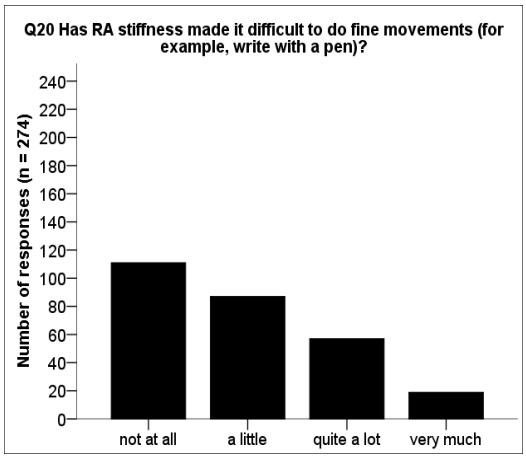


Figure 8.22: Frequency and distribution graph

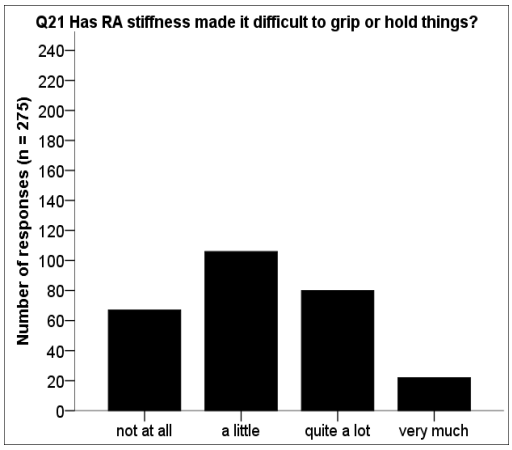


Figure 8.23: Frequency and distribution graph

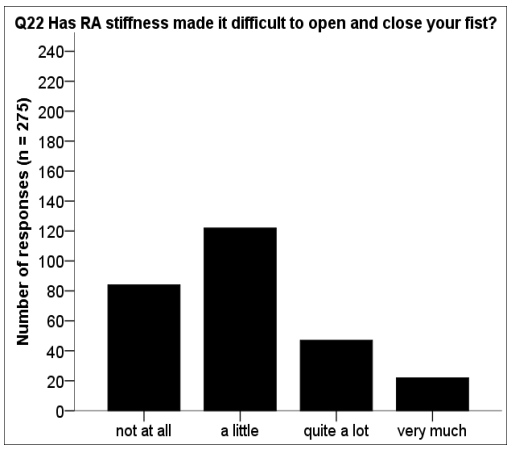


Figure 8.24: Frequency and distribution graph

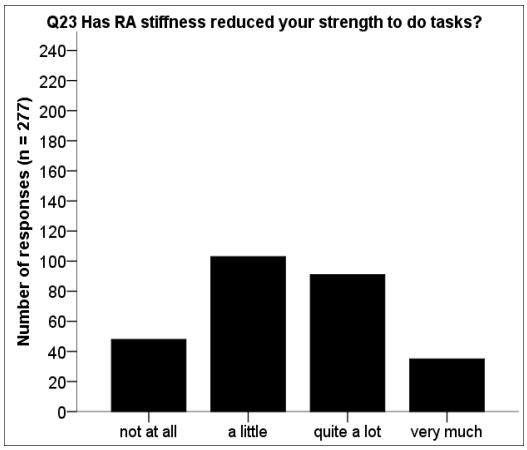


Figure 8.25: Frequency and distribution graph

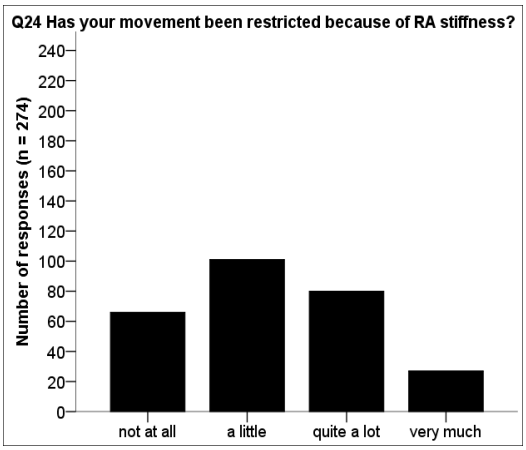


Figure 8.26: Frequency and distribution graph

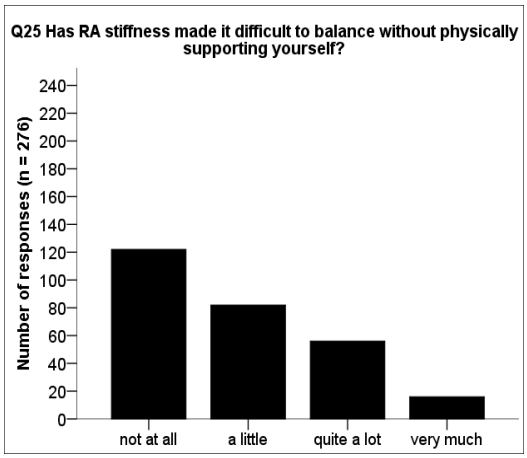


Figure 8.27: Frequency and distribution graph

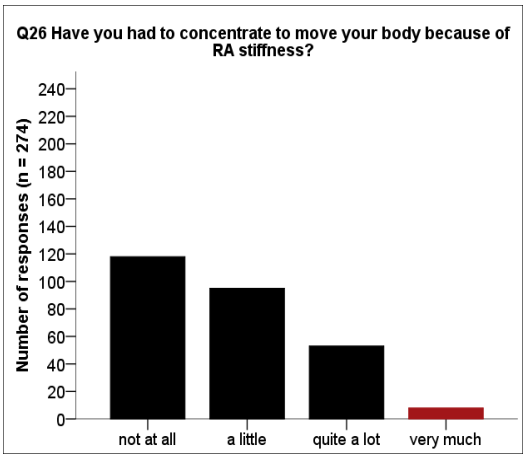


Figure 8.28: Frequency and distribution graph

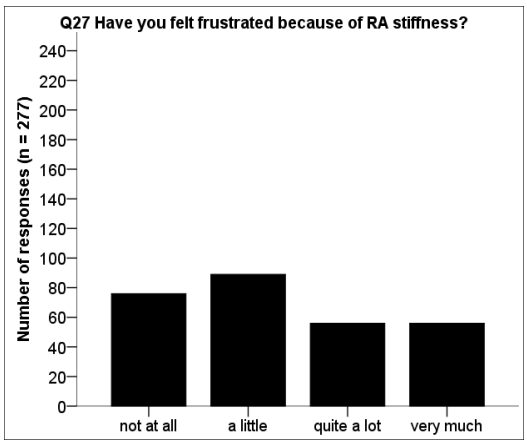


Figure 8.29: Frequency and distribution graph

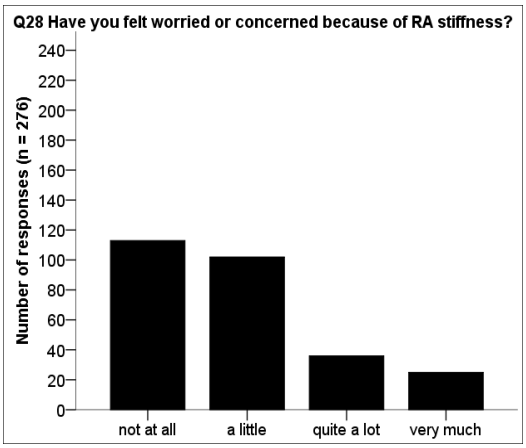


Figure 8.30: Frequency and distribution graph

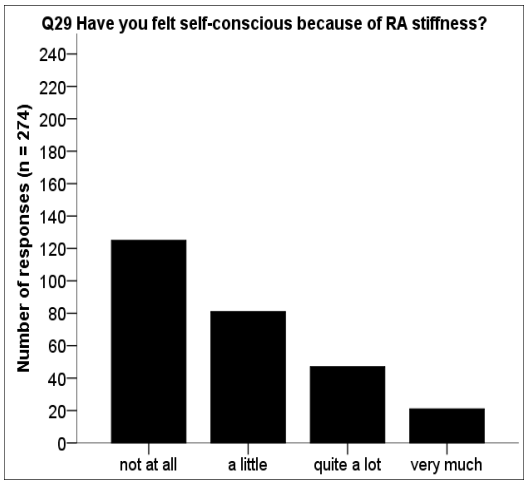


Figure 8.31: Frequency and distribution graph

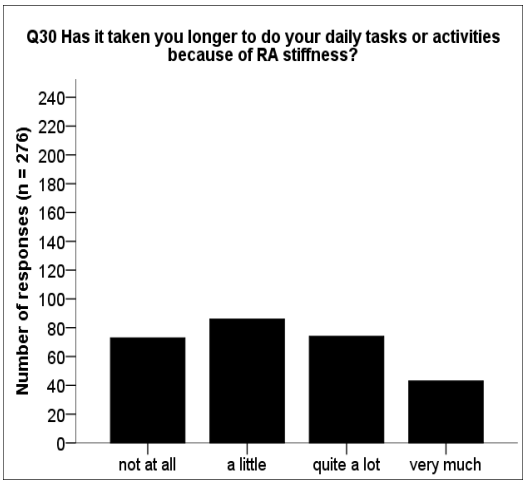


Figure 8.32: Frequency and distribution graph

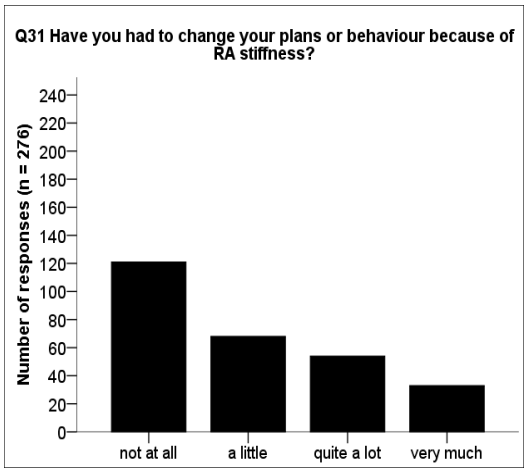


Figure 8.33: Frequency and distribution graph

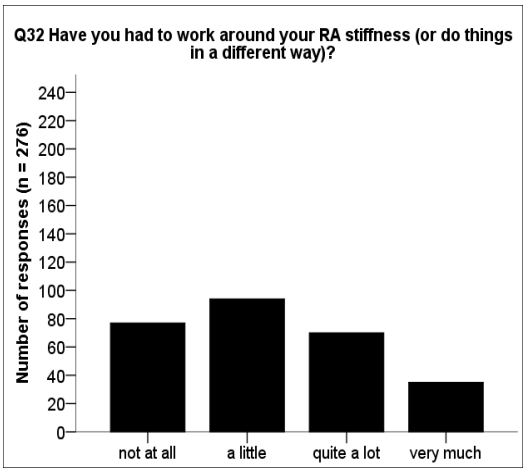


Figure 8.34: Frequency and distribution graph

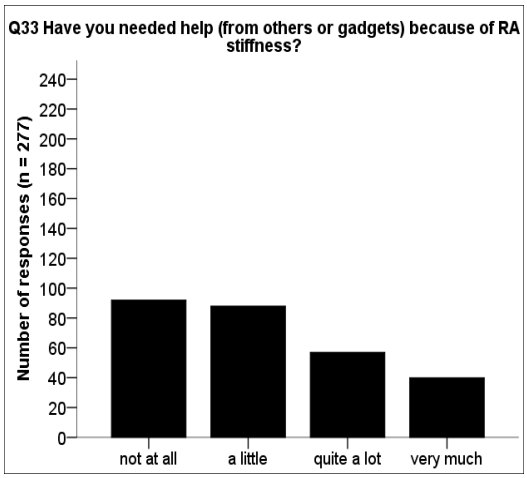


Figure 8.35: Frequency and distribution graph

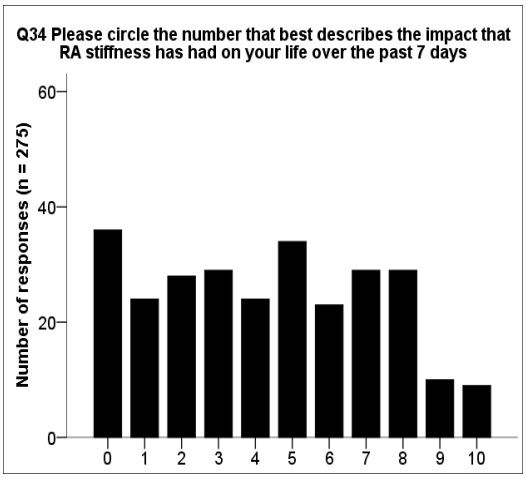


Figure 8.36: Frequency and distribution graph

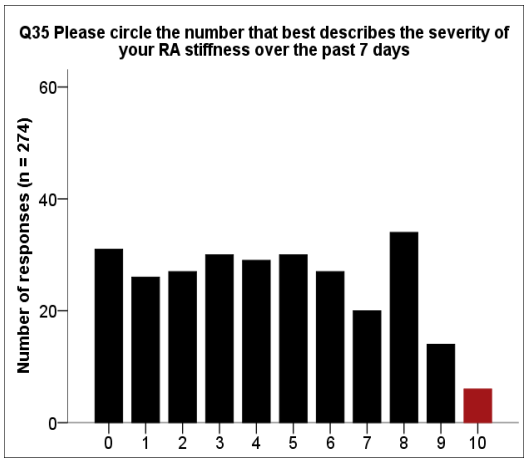


Figure 8.37: Frequency and distribution graph

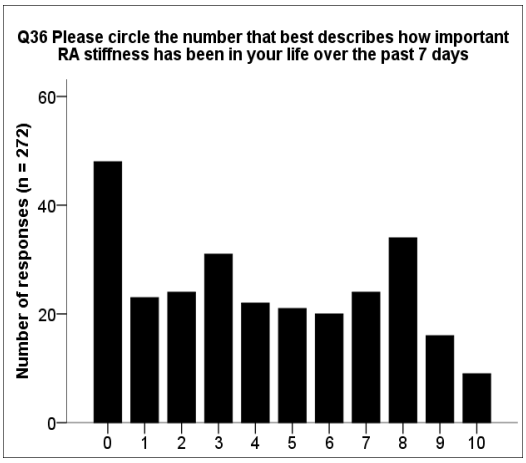


Figure 8.38: Frequency and distribution graph

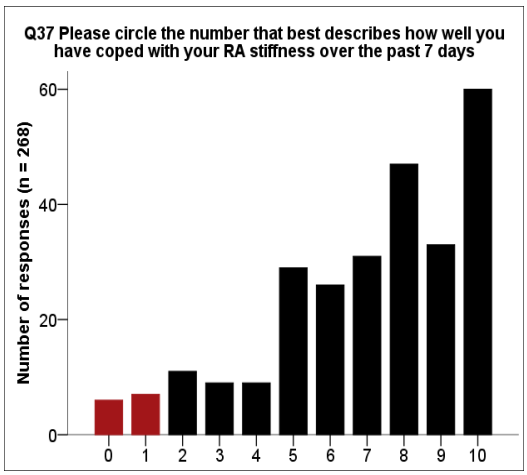


Figure 8.39: Frequency and distribution graph

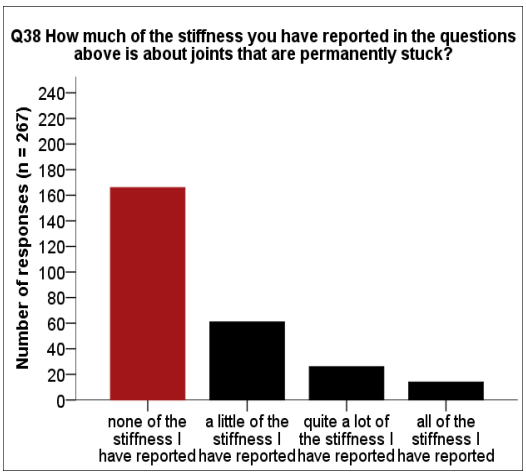


Figure 8.40: Frequency and distribution graph

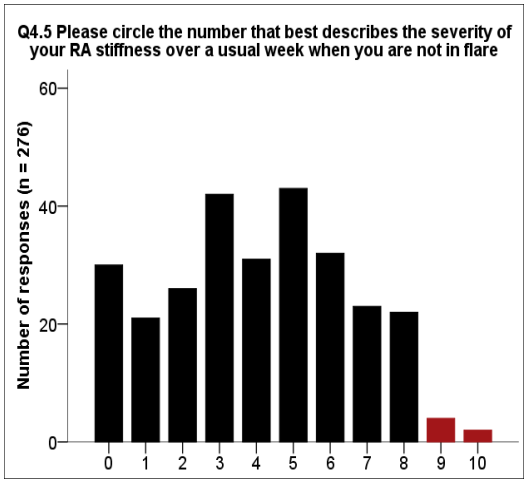


Figure 8.41: Frequency and distribution graph

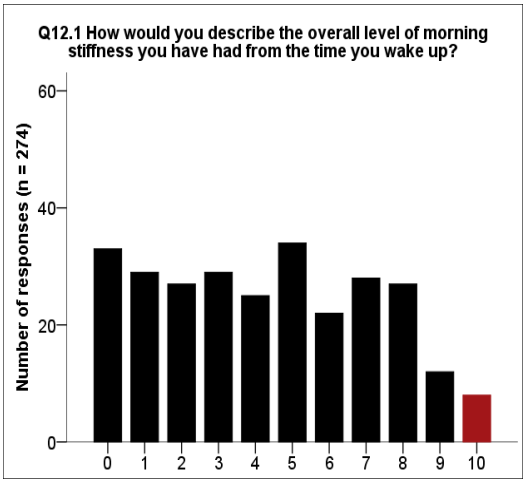


Figure 8.42: Frequency and distribution graph

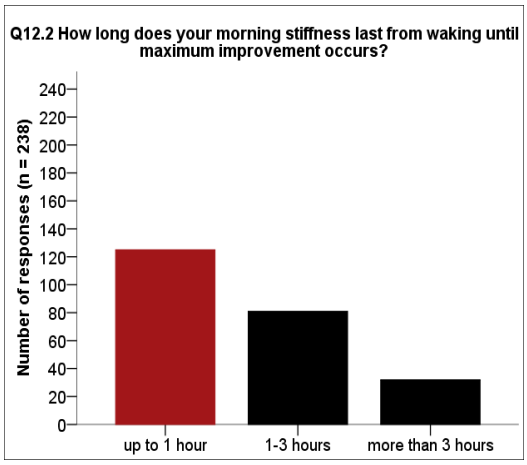


Figure 8.43: Frequency and distribution graph

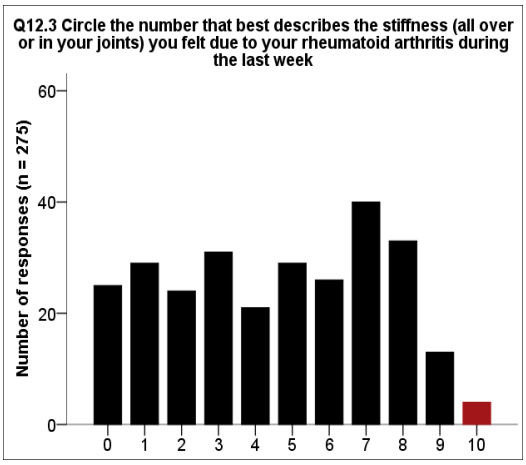


Figure 8.44: Frequency and distribution graph

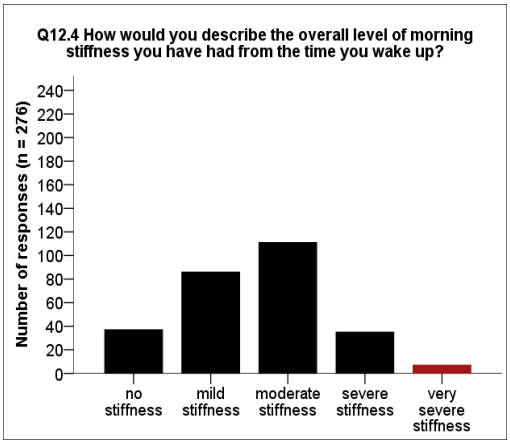


Figure 8.45: Frequency and distribution graph

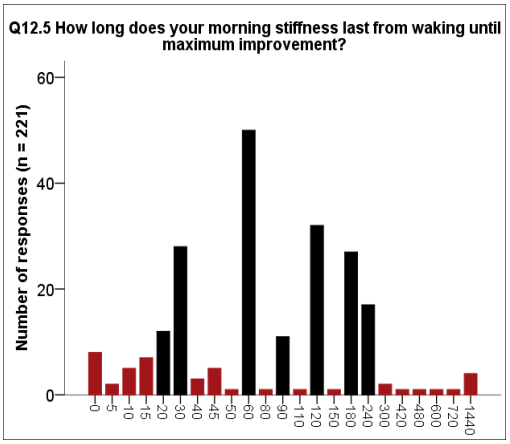


Figure 8.46: Frequency and distribution graph

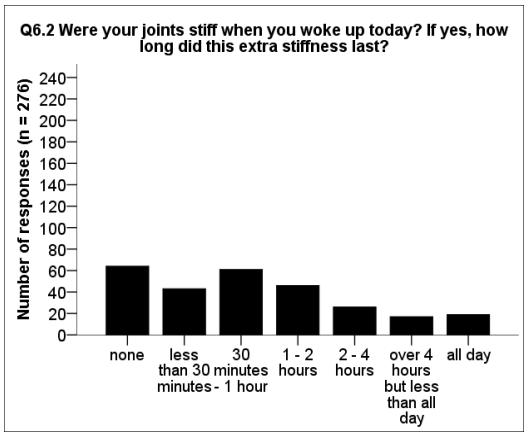


Figure 8.47: Frequency and distribution graph

8.4.4.1.3 Correlations

In the Pearson's correlation matrix of all variables (Appendix X) only items 'Do you have any joints that are permanently stuck?' and 'How long does your morning stiffness last from waking until maximum improvement occurs?' had very low correlations (r<0.30) in ≥36 (80%) of the 45 items. Three other items ('Over the past 7 days has your RA stiffness been as variable (coming and going) as usual for you?', 'Has RA stiffness made it difficult to chew?', 'How much of the stiffness you have reported in the questions above is about joints that are permanently stuck?') had very low correlations close to this threshold (>67%).

8.4.4.1.4 Consolidation of evidence regarding item suitability

Table 8.5 consolidates the information gathered from missing responses, distributions and frequencies, and correlations. Seven of the 45 draft items were identified as unsuitable at this stage: 'Please circle the number that best describes the severity of your RA stiffness over a usual week when you are not in a flare?', 'Do you have any joints that are permanently stuck?', 'Over the past 7 days has your RA stiffness been as variable (coming and going) as usual for you?', 'Has RA stiffness made it difficult to chew?', 'How much of the stiffness you have reported in the questions above is about joints that are permanently stuck?', 'How long does your morning stiffness last from waking until maximum improvement occurs?' (ordinal scale and minutes and hours). These are briefly reviewed below.

Although 'Please circle the number that best describes the severity of your RA stiffness over a usual week when you are not in a flare?' was developed as a draft stiffness item, on discussion it was decided that it should not be included in the questionnaire, therefore it was removed. Although 'Do you have any joints that are permanently stuck?' and 'How much of the stiffness you have reported in the questions above is about joints that are permanently stuck?' correlated moderately with each other (r=0.517, n=266, p<0.01), both demonstrated correlations <0.3 with most other items. The number of correlations <0.3 was over the 80% threshold for 'Do you have any joints that are permanently stuck?' and just under for 'How much of the stiffness you have reported in the questions above is about joints that are permanently stuck?', therefore both were considered unsuitable. 'Over the past 7 days has your RA stiffness been as variable (coming and going) as usual for you?' and 'Has RA stiffness made it difficult to chew?' had many but <80% correlations <0.3 with other items but had large percentages of responses in one response category, and small response ranges, and were removed. In contrast, while 'Over the past 7

days have you experienced RA stiffness all over?' and 'Over the past 7 days has your RA stiffness been different to usual for you?' had similar percentages of responses in one response category, they had considerably fewer correlations <0.3 with other items and therefore were retained. Both traditional items asking 'How long does your morning stiffness last from waking until maximum improvement occurs?' on different response scales (ordinal scale and minutes and hours) were removed because they had considerably higher amounts of missing data than other items and the concept of duration was captured in 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?'

Table 8.5: Draft items (n=45) and rationale for retention (n=38) or removal (n=7) based on response rate, distribution and correlations (assessing item suitability, Section 8.4.4.1)

Items	No. of missing values (n=277)	Pearson's correlations with other items <0.3 (n=45)	>50% of all responses in one category (%)	Categories with <8 responses (no. of categories, n=no. of responses)	Retain or remove
Draft stiffness items					
Draft item stuck joints A	2	42^	✓ (78.9%)	×	**Remove**
Draft item timing	6	6	×	×	Retain
Draft item in joints	1	4	✓ (63.8%)	✓ (1, n=7)	Retain
Draft item all over	0	17	✓ (83.4%)	×	Retain
Draft item different to usual	9	11	✓ (60.8%)	✓ (1, n=7)	Retain
Draft item variable	10	33	✓ (69.3%)	✓ (1, n=8)	**Remove**
Draft item after immobility	0	6	×	×	Retain
Draft item in body	4	8	✓ (50.5%)	✓ (1, n=6)	Retain
Draft item sleep	4	3	×	×	Retain
Draft item dress	4	5	×	×	Retain
Draft item wash	4	4	✓ (51.3%)	×	Retain
Draft item responsibilities	9	4	×	×	Retain
Draft item daily tasks	4	3	×	×	Retain
Draft item chew	3	30	✓ (89.4%)	✓ (2, n=6, n=3)	**Remove**
Draft item hobbies	8	5	×	×	Retain
Draft item get out of bed	1	6	×	✓ (1, n=7)	Retain
Draft item get up after sitting	2	6	×	×	Retain
Draft item effort	4	5	×	×	Retain
Draft item slower	4	4	×	×	Retain
Draft item fine movement	3	5	×	×	Retain
Draft item grip	2	6	×	×	Retain
Draft item open/close fist	2	6	×	×	Retain
Draft item strength	0	6	×	×	Retain
Draft item movement	3	5	×	×	Retain

Items	No. of missing values (n=277)	Pearson's correlations with other items <0.3 (n=45)	>50% of all responses in one category	Categories with <8 responses (no. of categories, n=no. of responses)	Retain or remove
Draft item balance	1	8	×	×	Retain
Draft item concentrate	3	5	×	✓ (1, n=8)	Retain
Draft item frustrated	0	4	×	×	Retain
Draft item worried	1	4	×	×	Retain
Draft item self-conscious	3	8	×	×	Retain
Draft item take longer	1	5	×	×	Retain
Draft item change plans	1	7	×	×	Retain
Draft item work around	1	4	×	×	Retain
Draft item need help	0	9	×	×	Retain
Draft item impact	2	2	×	×	Retain
Draft item severity	3	2	×	✓ (1, n=6)	Retain
Draft item importance	5	3	×	×	Retain
Draft item coped	9	22	×	✓ (2, n=6, n=7)	Retain
Draft item stuck joints B	10	35	✓ (62.2%)	×	**Remove**
Draft item response shift	1	5	×	✓ (2, n=4, n=2)	**Remove**
Traditional stiffness items					
Traditional item duration D	1	5	×	×	Retain
Traditional item severity A	3	3	×	✓ (1, n=8)	Retain
Traditional item duration E	39*	11	✓ (52.5%)	×	**Remove**
Traditional item severity G	2	4	×	✓ (1, n=4)	Retain
Traditional item severity C	1	3	×	✓ (1, n=7)	Retain
Traditional item duration F	56*	40^	×	✓ (16, n=<8 in all)	**Remove**

*=>5% missing values; ^>80% cut-off

8.4.4.2 Preliminary analysis

8.4.4.2.1 PCA

PCA on the 38 retained draft items gave a KMO measure of sampling adequacy of 0.98, which is defined as 'superb' (Kaiser, 1974, in Field, 2009). The diagonal elements of the anti-image matrices output for each of the KMO values for individual items were >0.94 and therefore above the recommended value (Field, 2009). Bartlett's test of sphericity was significant (X^2(703 degrees of freedom)=10892.48, p=<0.001) confirming that the correlations between items were sufficiently large to perform PCA (Pett, Lackey and Sullivan, 2003; Field, 2009).

8.4.4.2.2 NLPCA

NLPCA was performed three times on the 38 retained draft items specifying a 4-component solution and using different analysis levels giving the following VAF: nominal (73.28%), ordinal (73.09%) and numeric (71.19%). Only very slight improvement in VAF was seen by allowing non-linear (nominal or ordinal) transformations. Although an ordinal analysis level would be considered the most appropriate for these data, a numeric analysis level would provide greater stability, simpler interpretation, and less likelihood of results occurring by chance (Linting and van der Kooij, 2012).

8.4.4.3 Component extraction

PCA of the 38 retained draft items indicated a 4-component solution explaining 71.19% of the variance. NLPCA was performed specifying an ordinal analysis level and a 4-component solution which explained 73.09% of the variance.

8.4.4.4 Component rotation

PCA and NLPCA were performed on the 38 retained draft items first specifying orthogonal varimax rotation. Rotation clarified that each of the 4-components identified in component extraction were made up of a cluster of items with loadings of ≥0.40. The rotated solutions (Table 8.6) were very similar and most items loaded highest onto the same components across analyses. There were some minor differences including that items 'Have you had to work around your RA stiffness (or do things in a different way)?', 'Has it taken you longer to do your daily tasks or activities because of RA stiffness?', 'Has RA stiffness affected your sleep?', and 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' did not load highest onto the same components across

analyses. ‘Have you had to work around your RA stiffness (or do things in a different way)?’ and ‘Has it taken you longer to do your daily tasks or activities because of RA stiffness?’ had loadings ≥0.40 on both components 1 and 3 but the highest was different between analyses. ‘Has RA stiffness affected your sleep?’ loaded ≥0.40 on both component 1 and 4 in PCA but only on component 1 in NLPCA. ‘Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days’ failed to load ≥0.40 on any component in PCA but loaded ≥0.40 on component 2 in NLPCA.

PCA and NLPCA were also performed specifying oblique promax rotation. The rotated pattern and structure matrices (Appendix Y) were similar to the orthogonal varimax rotation, with items loading on similar components. As varimax rotation is more commonly used and is easier to interpret as it produces a single matrix (Pett, Lackey and Sullivan, 2003), this approach was preferred.

Table 8.6: Eigenvalues, VAF, rotated component loadings, and communalities across PCA and NLPCA solutions

Items	Rotated component loadings								Communalities	
	PCA				NLPCA*				PCA	NLPCA*
	1	2	3	4	1	2	3	4		
Draft item wash	**.791**	.219	.199	.274	**.823**	.234	.134	.184	.774	.769
Draft item dress	**.742**	.281	.202	.345	**.790**	.295	.114	.281	.789	.790
Draft item grip	**.688**	.362	.313	.139	**.683**	.394	.305	.047	.722	.718
Draft item responsibilities	**.679**	.372	.337	.219	**.734**	.327	.311	.167	.761	.765
Draft item balance	**.669**	.161	.328	.182	**.723**	.177	.219	.131	.614	.621
Draft item daily tasks	**.668**	.437	.296	.263	**.728**	.381	.271	.199	.795	.785
Draft item fine movement	**.659**	.341	.273	.160	**.703**	.299	.234	.103	.651	.648
Draft item get out of bed	**.647**	.182	.222	.472	**.679**	.262	.128	.413	.724	.709
Draft item need help	**.611**	.226	.485	.186	**.671**	.205	.435	.119	.695	.710
Draft item movement	**.608**	.358	.465	.241	**.699**	.352	.413	.180	.773	.774
Draft item strength	**.606**	.344	.466	.130	**.622**	.360	.456	.062	.720	.736
Draft item concentrate	**.583**	.207	.442	.242	**.649**	.229	.335	.209	.636	.631
Draft item open/close fist	**.581**	.448	.248	-.040	**.590**	.397	.272	-.040	.601	.587
Draft item effort	**.564**	.435	.456	.282	**.632**	.411	.426	.221	.795	.794
Draft item hobbies	**.562**	.301	.471	.315	**.640**	.307	.409	.264	.734	.728
Draft item slower	**.549**	.464	.486	.194	**.618**	.421	.479	.107	.791	.795
Draft item get up after sitting	**.543**	.255	.290	.490	**.574**	.376	.228	.398	.984	.664
Draft item severity	.368	**.690**	.401	.347	.353	**.774**	.383	.170	.893	.908
Draft item different to usual	.121	**.689**	.047	.124	.181	**.593**	-.143	.103	.507	.510
Traditional item severity G	.380	**.677**	.388	.338	.354	**.802**	.341	.157	.869	.961
Traditional item severity C	.395	**.672**	.260	.345	.336	**.778**	.288	.157	.794	.831
Draft item impact	.398	**.628**	.477	.338	.423	**.689**	.445	.192	.895	.896
Traditional item duration D	.245	**.626**	.320	.231	.331	**.653**	.294	.180	.608	.672
Traditional item severity A	.433	**.625**	.312	.381	.361	**.797**	.318	.155	.821	.890

Items	Rotated component loadings								Communalities	
	PCA				NLPCA*				PCA	NLPCA*
	1	2	3	4	1	2	3	4		
Draft item importance	.354	**.598**	.529	.326	.396	**.614**	.537	**.198**	.870	.861
Draft item timing	.243	**.598**	.529	.326	.169	**.796**	.159	.210	.585	.704
Draft item in joints	.341	**.598**	.529	.326	.199	**.830**	.044	.095	.620	.759
Draft item coped	-.215	-.362	-.215	.026	-.225	**-.537**	-.397	-.005	.224	.459
Draft item worried	.228	.234	**.771**	.221	.351	.205	**.714**	.250	.750	.741
Draft item self-conscious	.323	.155	**.743**	.252	.419	.160	**.670**	.235	.745	.702
Draft item frustrated	.412	.313	**.679**	.277	.474	.351	**.639**	.219	.806	.809
Draft item change plans	.475	.225	**.678**	.151	.577	.164	**.609**	.219	.759	.766
Draft item work around	.572	.243	**.602**	.187	**.636**	.241	.550	.144	.784	.792
Draft item take longer	.569	.373	**.577**	.211	**.651**	.341	.535	.151	.840	.859
Draft item all over	.066	.050	.174	**.764**	.128	.104	.129	**.805**	.620	.696
Draft item in body	.178	.248	.229	**.672**	.246	.231	.190	**.726**	.597	.679
Draft item after immobility	.319	.331	.224	**.599**	.356	.464	.195	**.485**	.621	.621
Draft item sleep	.443	.344	.267	**.451**	**.505**	.376	.225	.371	.590	.583
Initial eigenvalues	**22.98**	**1.80**	**1.18**	**1.10**	**22.87**	**2.47**	**1.35**	**1.08**		
Initial VAF	**60.47%**	**4.73%**	**3.09%**	**2.90%**	**60.19%**	**6.50%**	**3.56%**	**2.83%**		
Rotated eigenvalues	**9.62**	**6.64**	**6.40**	**4.40**	**11.12**	**8.35**	**5.49**	**2.82**		
Rotated VAF	**25.31%**	**17.47%**	**16.84%**	**11.58%**	**29.26%**	**21.97%**	**14.45%**	**7.41%**		
Total VAF	**71.19%**				**73.09%**					

NB **bold loadings**=highest loading; underlined loadings=other loadings ≥0.4; *Specifying an ordinal analysis level

8.4.4.5 Comparing the two statistical approaches

The output from PCA was compared to the output generated from NLPCA specifying an ordinal analysis level which was considered most appropriate for these data. Four criteria were used to compare PCA and NLPCA; 1) eigenvalues and VAF; 2) VAF across components 3) component loadings; and 4) communalities (Table 8.6).

As expected, initial eigenvalues for both solutions were large for the first component (22.98 and 22.87 for PCA and NLPCA respectively) and much smaller for the following three components. Following rotation, the eigenvalues and VAF in both solutions demonstrated a more even spread across components. Total VAF was 71.19% and 73.09% for PCA and NLPCA respectively. PCA demonstrated slightly more even VAF across components which was considered preferable from the perspective of conceptual importance and stability and robustness for further testing. Component loadings were similar, despite small differences in the loadings of some items between solutions, items generally loaded on the same components. Similarly, communalities were not identical but were consistent across methods (e.g. items with low communalities were generally low in both PCA and NLPCA).

Benefits of PCA include increased stability, simpler interpretation, and less risk of occurring by chance (Linting and van der Kooij, 2012). A further advantage of PCA is that calculations are generated from the correlation matrix whereas in NLPCA calculations are generated directly from the data (Linting and van der Kooij, 2012). As a result, in NLPCA, the solution can be "over fitted" to the particular dataset in which a solution is developed, which may influence the stability of the solution when tested in other populations. PCA also provides all aspects of a rotated analysis as part of one programme and has the additional benefit of familiarity in the literature which may also improve the accessibility of the results. Although in theory PCA is less appropriate for the data available here, the comparison between PCA and NLPCA shows similarity between the outputs regardless. Overall, given the similarity between the PCA and NLPCA outputs and the advantages of PCA, PCA was taken forward to perform further testing to develop a new RA stiffness PROM (Chapter 9).

8.5 Discussion

8.5.1 Postal survey

The survey sample reflected the characteristics of patients in other recent survey research (e.g. Wilson *et al,* 2015; Hammond *et al,* 2015) and included a range of

participants from disease related and sociodemographic perspectives. The response rate of 43.1% was slightly lower than that found in the aforementioned studies, but similar to other questionnaire development studies (e.g. Goodacre *et al,* 2007). However, response rate may have been affected by the data collection approach where information including consent was collected altogether.

A strength of the survey is the collection of non-responder data. Non-response bias occurs when those who do not participate are different from those who do participate in respect to the topic under investigation (Silman and Macfarlane, 2002). Collection of non-responder data in this study enabled demonstration that the frequencies and percentages of gender, age, and social deprivation were similar in responders and non-responders. A limitation of the survey was that the questionnaire pack was not piloted prior to use. An implication of this could have been reducing the amount of missing data. However, the draft stiffness items and traditional unvalidated stiffness items had been tested with patients in cognitive interviews (Chapter 7), and the questionnaire pack was developed with patient partner (GB) involvement. Despite this, a small pilot of the questionnaire would likely have enhanced data completion and response rates, provided valuable information and been consistent with recommendations (e.g. Silman and Macfarlane, 2002).

A small number of participants (n=6, 2.4%) reported that they had other conditions in addition to RA, where stiffness was also a feature (Section 8.4.2.2.3). That patients with other rheumatic conditions were not excluded from this study may be perceived as a limitation. However, this study did not exclude patients with other conditions so as to not limit the sample and reduce the questionnaires application clinically and in research. For example, OA is very common in the general population, with recent figures indicating that one third of individuals >45 years have sought treatment for OA (Arthritis Research UK, 2013b). Such figures indicate crossover with the age of an RA population and could have reduced the sample considerably. In addition, the final questionnaire will need to be robust enough to be used in the target population, which will likely contain individuals with other conditions such as OA. Interestingly, the numbers reported in this study were much lower than expected based on the OA figures reported by Arthritis Research UK (2013b), but were consistent with those found in another recent survey study where 8.2% of the sample (n=413) reported having other rheumatic conditions, specifically 19 (5%) reported OA (Wilson, 2016). It must be acknowledged however that both studies were based on patient self-reported information. The accuracy of patient self-reports of rheumatic conditions

have been challenged by some authors (e.g. Kvien *et al,* 1996) yet reported as accurate by others (e.g. Barlow, Turner and Wright, 1998). Overall, further consideration of the study population could be explored in more detail in future validation studies and is certainly relevant in future research directions relating to the universal or specific nature of stiffness assessment tools (Section 2.4.2 and Chapter 10). Regardless, the careful development of the draft items, for example the inclusion of RA stiffness in each stem question, will enhance clarity and specificity, and hopefully aid the robustness of items in such situations.

8.5.2 Analysis approach

Although not originally considered within the plan for this research, the recognition of and exploration into the debate surrounding the methodological appropriateness of different analytical approaches has turned into a strength of this work. As discussed earlier (Section 8.1.3), when performing PCA, current recommendations suggest the use of alternative approaches to producing Pearson's correlations, such as polychoric correlations, when working with data that are not continuous (Streiner and Norman, 2008). Despite this, these recommendations are not routinely employed in practice. A recent review of EFA (and PCA) in nursing studies identified 54 analyses in 28 papers, all of which were performed using methods suited to continuous data despite all being based on ordinal (91%) or nominal (9%) data (Gaskin and Happell, 2014). The authors suggest that this may be expected given the limited guidance regarding performing factor analysis with ordinal and nominal data in comparison to guidelines for continuous data. In combination with other restrictions including the poor accessibility of programmes for these purposes (Baglin, 2014) and limitations of their use (e.g. Lorenzo-Seva and Ferrando, 2014), this may provide some explanation for the poor uptake of recommendations in practice. Given the lack of guidance and difficulties regarding implementation of non-traditional approaches to PCA, it is rare to see comparisons of different methodological approaches, especially in practical rather than theoretical examples. This chapter has described a novel comparison of two analytical approaches; the traditional approach suited to the analysis of continuous data (PCA), and an approach suited to the analysis of non-continuous data (NLPCA). It has provided evidence of consistent results across analyses and enabled evidence based decision making regarding PCA as acceptable for further analysis, a decision which would otherwise be inconsistent with current recommendations (e.g. Streiner and Norman, 2008).

Another consideration is the use of CTT (Section 8.1.1) as a basis for PROM development. Although CTT is broadly applicable and well used (Hambleton and Jones, 1993), the newer IRT is considered by some to be a better alternative as it overcomes the limitations of CTT (Streiner and Norman, 2008). The advantages of IRT models include true interval properties, more precise estimation of measurement error, and test-free measurement (comparison of subjects regardless of the items completed) (Streiner and Norman, 2008). Despite these advantages, IRT models do not directly improve item wording or construction (DeVellis, 2012), therefore the use of CTT does not detract from the careful development of items in this study. Within the PROM development literature, recommendations for the use of IRT models are now common place. For example, the FDA recently stated that they recognise the benefits of modern approaches to scale development such as IRT methods over more traditional approaches. However, they also recommend that sponsors include assessment of both traditional and novel approaches to demonstrate links between methods (Patrick *et al,* 2007). This dual-method recommendation is reflected in the theoretical literature where rather than IRT methods replacing more traditional approaches, it is recognised that both approaches have advantages (DeVellis, 2012). IRT and CTT have been described as "complementary approaches" (de Champlain, 2010, p.117) which "should be integrated in a comprehensive approach to measurement issues" (Embretson and Hershberger, 1999, p.252). Given the above considerations, any stiffness PROM developed from this research would benefit from further exploration in future research using newer IRT models such as Rasch analysis (e.g. Tennant and Conaghan, 2007).

8.6 Conclusion

This chapter has described the demographics of participants involved in the postal survey and found the sample to be similar to other recent survey studies and representative of the RA population. The comparison of PCA and NLPCA identified considerable similarities in the outputs which facilitated the decision that PCA was the most appropriate method to further analyse the dataset given its improved stability, familiarity, and practicality. Of the 45 draft stiffness items, seven were established as being unsuitable for further analysis and were therefore removed. The remaining 38 draft stiffness items will now be taken forward to Chapter 9 for further development and testing using PCA to identify the smallest and most effective item structure for a new RA stiffness PROM.

Chapter 9: Developing and testing the structure of an RA stiffness PROM (Study 4, part 2)

The previous chapter reported on the survey methods and sample demographics, and addressed which analytical approach would be most appropriate for the development of a new RA stiffness PROM. PCA was identified as the most appropriate method. This chapter now describes the use of PCA to develop the smallest and most effective item structure utilising the remaining 38 draft items, and then describes preliminary testing of the resulting RA stiffness PROM.

9.1 Background

9.1.1 Developing the provisional RA stiffness PROM

In using PCA to develop a provisional RA stiffness PROM, the smallest item structure was sought for feasibility (Boers *et al,* 1998) in clinical and research environments. This required attention to internal consistency which is concerned with the homogeneity of the items that make up an instrument (Field, 2009; DeVellis, 2012) (Section 2.3.1.3). To investigate the coherence of the whole item structure and each individual component, and to identify items that contribute least, Cronbach's alpha was performed in combination with PCA. This was supplemented by performing successive PCAs on each individual component to investigate whether items constitute a single component or a number of smaller components (Pett, Lackey and Sullivan, 2003). Consideration of statistical criteria, theoretical appropriateness and simplicity are recommended during solution refinement in PCA (Nunnally and Bernstein, 1994; Pett, Lackey and Sullivan, 2003). Therefore, all decisions regarding item reduction and the final structure of the provisional RA stiffness PROM were tempered with information from statistical tests and expert (clinical and patient) judgement.

9.1.2 Concepts of measurement

As appropriate measurement properties are required for PROMs to be useful (Terwee *et al,* 2007), following the development of the provisional RA stiffness PROM, preliminary testing of some measurement properties was performed. A valid measure (one which measures what it intends to measure (SACMOT, 2002; Frost *et al,* 2007)) would have appropriate relationships with other measures of RA (construct validity, Section 2.3.1.1). Relationships have been highlighted in the qualitative literature (e.g. Lineker *et al,* 1999; Orbai *et al,* 2014; Halls *et al,* 2015) between stiffness and disability, pain, fatigue, and flare. Therefore, moderate correlations would be expected

between the provisional RA stiffness PROM and measures capturing these constructs. Stiffness is also considered an indicator of inflammatory activity in RA (e.g. Lansbury, 1956; Scott, 1986; Hazes *et al,* 1994; Soubrier *et al,* 2006). Therefore, moderate to strong correlations would be expected between the provisional RA stiffness PROM and measures capturing disease activity. Correlation cut-offs were as defined previously (Section 2.3.1.1.1) and testing of expected relationships used data that were not involved in decisions regarding PROM development.

9.2 Aims and objectives

The overall aim of this study was to develop and test the structure of a provisional RA stiffness PROM. The specific objectives of this chapter were:

- From the 38 provisional stiffness items, to identify the smallest and most effective item structure to form a new provisional RA stiffness PROM
- To test the validity of the provisional RA stiffness PROM by investigating its correlations with other measures of RA
- To provide recommendations on the most appropriate tool to use to assesses stiffness in research and clinical situations

9.3 Methods

The methods for the survey in Study 4 were described in Chapter 8 (Section 8.3).

9.4 Analysis and results

Analysis stages 1 and 2 were described in Chapter 8 (Section 8.4). Stage 3 analysis involved statistical analysis guided by expert (clinical and patient) judgement to reduce the number of items into the smallest and most effective item structure (item reduction and structure development). Stage 4 involved testing of the provisional RA stiffness PROM using comparisons with other measures of RA (validity testing). The description of each analysis stage is followed directly by its results.

9.4.1 Stage 3 analysis: Item reduction and structure development

Item reduction and structure development was performed in seven steps: 1) explore the whole structure; 2) explore each component individually; 3) test item removal; 4) decide on the final component structure; 5) bootstrapping; 6) scoring; 7) formatting.

9.4.1.1 Explore the whole structure

In Chapter 8, loadings ≥0.40 were defined as high or meaningful. However, the aim of stage 3 analysis was to refine the solution therefore, stricter item loading criteria were adopted. Guidelines suggest that 0.45='fair', 0.55='good', 0.63='very good', 0.71='excellent', and it is recommended that the more 'very good' or 'excellent' loadings the better (Comrey and Lee, 1992). Therefore, 'good' item loadings (≥0.55) were targeted to seek the optimal model.

Even after rotation, it is common for PCA solutions to include items with high loadings on several components, and items with weak loadings on all components. There is conflicting advice about what to do with such items (Pett, Lackey and Sullivan, 2003). Some suggest removing items with high loadings on many components to reduce difficulties with interpretation (Kline, 2000). Others recommend careful consideration of their placement (Hair *et al,* 1995), taking into account the conceptual relationship between item and component (Pett, Lackey and Sullivan, 2003). Given the importance of the conceptual perspective of the item structure and to retain as much useful information as possible, the latter approach was implemented.

Consistent with the idea that PCA aims to identify groups of related items (Pallant, 2010), it is also important that those items are homogenous (DeVellis, 2012) and 'hang together' well. This can be assessed by Cronbach's alpha coefficient which evaluates internal consistency (Pett, Lackey and Sullivan, 2003). In questionnaires with different components, it is also recommended that the internal consistency of each component is investigated (Field, 2009). Therefore Cronbach's alpha coefficient was explored in the whole structure and each individual component. The output from these analyses provide information about the effect that the deletion of each item would have on the internal consistency of the scale (Pallant, 2010), and aids decisions about item removal. The correlation of an item with the component having removed that item (corrected item total correlation) will be strong and positive if internal consistency is demonstrated (Pett, Lackey and Sullivan, 2003). This can be explored further using the correlation matrix of all items within the component. The correlation matrix generated from PCA rather than from internal consistency analysis was used as PCA can be performed pairwise while internal consistency analysis can only be performed listwise, thus utilising more available data.

9.4.1.2 Explore each component individually

PCA was performed on each individual component to establish whether items constituted a single component or a number of smaller components (Pett, Lackey and Sullivan, 2003). Investigation into item loadings and internal consistency was also performed as described above.

9.4.1.3 Test item removal

Items that do not contribute to the homogeneity of the scale can be removed without loss of information, and it is recommended that items with low loadings are removed and analyses rerun without them (Hair *et al,* 1995). A series of rotated PCA were performed alongside internal consistency analyses. During each round, items with the lowest component loading or which, if removed, would increase the overall alpha were considered for removal. Decisions about which item to remove (or replace) at each iteration were based on the output of the PCA at that iteration, information identified above (Sections 9.4.1.1 and 9.4.1.2), and in previous examination of items (Section 8.4.4). Following removal of an item, PCA was repeated and the resulting solution was reviewed. This process was continued until an appropriate solution, from both statistical and conceptual perspectives, was identified (Section 9.4.1.4).

9.4.1.4 Decide on the final component structure

Decisions made during item removal (Section 9.4.1.3) were based on statistical information (e.g. component loadings). It was also important to identify whether this could be improved upon by consideration from a conceptual perspective. Therefore, following item removal, the component structure was discussed with the supervisory team. Discussions considered whether the content of the individual components made sense conceptually and corresponded with results from the qualitative studies (Chapters 4, 5 and 7), informed how components should be labelled to reflect their content, and influenced further item removal testing for example by identifying items with similar wording.

9.4.1.5 Bootstrapping

The provisional RA stiffness PROM structure was then tested for stability and robustness using bootstrapping. Bootstrapping involved performing 20 repeated PCA analyses on randomly selected subsets of data (approximately 50% of the whole sample per subset) to establish whether the component structure was retained.

9.4.1.6 Scoring

Consideration was given to how to combine the responses to each item in a meaningful way, or score the scale (Streiner and Norman, 2008). Scores can be generated using simple summation or based on factor scores generated in PCA. The simple summation approach involves adding together the response to each item as assigned during coding. Factor scores are generated from standardised participant scores, which are weighted by a generated coefficient and then summed across items (Pett, Lackey and Sullivan, 2003). Factor scores can be performed in three ways; regression method, Bartlett method, Anderson-Rubin method. Although recommendations regarding the appropriate approach vary it has been suggested that for most studies, the three approaches will produce similar factor scores (Kim and Mueller, 1978). As the regression method is commonly employed (Pett, Lackey and Sullivan, 2003), it was used in this study. To investigate the relationship between scores generated from simple summation and from factor scores, correlations and scatterplots were produced and compared.

9.4.1.7 Formatting

Finally, discussion with the supervisory team and patient research partners developed names and other formatting aspects such as the layout, font, and order of items in the provisional RA stiffness PROM.

9.4.2 Stage 3 results: Item reduction and structure development

9.4.2.1 Explore the whole structure

When reviewing the whole item structure, component loadings could be generally described as ‘good’ (≥0.55). However, three items (‘Has RA stiffness affected your sleep?’, ‘Has RA stiffness made it difficult to get up after sitting for a while?’ and ‘Has RA stiffness made you slower (for example, unable to do things quickly)?’) demonstrated loadings just below this threshold and one item (‘Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days’) demonstrated weak loadings on all components (Table 9.5). The strongest loading for ‘Has RA stiffness made it difficult to get up after sitting for a while?’ (0.543) and ‘Has RA stiffness made you slower (for example, unable to do things quickly)?’ (0.549) were marginally below the ‘good’ threshold, and ‘Has RA stiffness affected your sleep?’ (0.451) just met the criteria of a ‘fair’ loading (>0.45). ‘Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days’ had a weak loading on all components.

The whole 38-item 4-component structure indicated a homogenous set of items (Cronbach's alpha=0.961). The removal of 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' would result in a slight improvement to internal consistency (Cronbach's alpha=0.971).

9.4.2.2 Explore each component individually

PCA was performed on each of the four individual components to establish whether they were made up of more than one component and to investigate internal consistency.

9.4.2.2.1 Component 1

A single component was identified and all item loadings were 'excellent' (Table 9.1). The Cronbach's alpha coefficient for the 17-items was 0.973 which could not be improved by item removal. Generally, correlations between items were moderate or strong, apart from between 'Has RA stiffness made it difficult to open and close your fist?' and 'Has RA stiffness made it difficult to get out of bed?', 'Has RA stiffness made it difficult to get up after sitting for a while?' and 'Has RA stiffness made it difficult to balance without physically supporting yourself?' which were weak (r=≤0.46). 'Has RA stiffness made it difficult to open and close your fist?' also demonstrated the lowest component loading and Cronbach's alpha coefficient. From a conceptual perspective, component 1 appeared to contain items capturing impact on physical tasks and daily life activities, as such this was initially labelled as 'Physical'.

Table 9.1: Item loadings and Cronbach's alpha coefficient for component 1 ("Physical")

Item	Loading	Cronbach's alpha coefficient
Draft item effort	.884	.872
Draft item daily tasks	.884	.868
Draft item movement	.879	.867
Draft item slower	.871	.863
Draft item responsibilities	.868	.846
Draft item dress	.855	.835
Draft item hobbies	.852	.831
Draft item grip	.841	.841
Draft item strength	.841	.836
Draft item wash	.834	.816
Draft item need help	.813	.797
Draft item fine movement	.796	.779
Draft item get out of bed	.794	.767
Draft item concentrate	.791	.772
Draft item get up after sitting	.783	.773
Draft item balance	.768	.746
Draft item open/close fist	.715	.698

9.4.2.2.2 Component 2

A single component was identified and item loadings were generally 'excellent', although 'Over the past 7 days has your RA stiffness been different to usual for you?' was 'good', and 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' was 'weak' (Table 9.2). The weak loading of 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' (-0.414) in component 2 was consistent with the poor loading identified for that item within the whole item structure (-0.362). The overall Cronbach's alpha coefficient for the 11-item component was 0.926. If 'Have you experienced RA stiffness in your joints over the past 7 days?' or 'Over the past 7 days has your RA stiffness been different to usual for you?' were removed the overall Cronbach's alpha coefficient would increase slightly to 0.929. If 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' was removed the overall Cronbach's alpha coefficient would increase to 0.935. Generally, correlations between items were moderate or strong, apart from between 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' and 'Over the past 7 days has your RA stiffness been different to usual for you?' and all other items which were weak. From a conceptual perspective, this component appeared to contain items capturing stiffness severity and broad stiffness impact, as such this was initially labelled as 'Severity'.

Table 9.2: Item loadings and Cronbach's alpha coefficient for component 2 ("Severity")

Item	Loading	Cronbach's alpha coefficient
Draft item severity	.955	.935
Traditional item severity G	.947	.924
Draft item impact	.942	.924
Traditional item severity A	.924	.896
Draft item importance	.921	.908
Traditional item severity C	.901	.864
Traditional item duration D	.769	.684
Draft item in joints	.728	.618
Draft item timing	.700	.600
Draft item different to usual	.587	.543
Draft item coping	-.414	.404

9.4.2.2.3 Component 3

A single component was identified and all item loadings were 'excellent' (Table 9.3). The overall Cronbach's alpha coefficient for the 6-item component was 0.941 which could not be improved by item removal. Correlations between all items were moderate or strong. Conceptually this component appeared to capture psychosocial impact, including emotional aspects and daily management, and was initially labelled as 'Psychosocial'.

Table 9.3: Item loadings and Cronbach's alpha coefficient for component 3 ("Psychosocial")

Item	Loading	Cronbach's alpha coefficient
Draft item frustrated	.903	.856
Draft item work around	.902	.853
Draft item ADLs take longer	.897	.847
Draft item change plans	.888	.835
Draft item self-conscious	.851	.788
Draft item worried	.833	.763

9.4.2.2.4 Component 4

A single component was identified and all item loadings were 'excellent' (Table 9.4). The overall Cronbach's alpha coefficient for the 4-item component was 0.761 which despite being 'acceptable', was lower than that of other components. Removal of any item failed to improve the overall Cronbach's alpha coefficient, and correlation between items were weak or moderate. From a conceptual perspective, this component was not as clear as other components as it contained items capturing stiffness location but also items capturing impact and severity. Initially this was

labelled as ‘Location’ but these results raise questions regarding this component in the model.

Table 9.4: Item loadings and Cronbach’s alpha coefficient for component 4 (“Location”)

Item	Loading	Cronbach’s alpha coefficient
Draft item in body	.819	.625
Draft item after immobility	.790	.613
Draft item sleep	.778	.595
Draft item all over	.733	.533

9.4.2.3 Test item removal

Box 9.1 describes each round of item removal including the rotated PCA output (Tables 9.5-9.25), justification for the item removed at each round and internal consistency analyses. As the item loadings changed and components moved position during the process, components are referred to by their initial label rather than by their number, and are distinguished by colour (as per the key). Only loadings >0.40 were reported for clarity, unless important for description. For all items, the highest loading has been illustrated in bold. For items which loaded on more than one component, underlining is used to illustrate all other loadings >0.40.

Box 9.1: Rotated PCA and internal consistency analyses for item removal

Description of item removal

The first rotated PCA (Table 9.5) represents the 38-item, 4-component solution with a Cronbach's alpha coefficient of 0.961 (Chapter 8). Here 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' (Draft item coped) was considered for removal as it had the lowest component loading in the PCA. 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' had been identified earlier as falling under the weak component loading threshold (<0.40) (Section 9.4.2.1) and performing poorly within examination of individual components (Section 9.4.2.2). It had also been identified in Chapter 8 as having a number (n=22) of poor correlations (<0.3) with other items (Section 8.5.2.1.3). When considering internal consistency, the removal of this item would slightly improve the overall Cronbach's alpha coefficient (to 0.971). Therefore, 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' was removed (Table 9.6).

Rotated PCA output

Table 9.5: Rotated PCA 1

Item	1	2	3	4
Draft item wash	**.791**			
Draft item dress	**.742**			
Draft item grip	**.688**			
Draft item responsibilities	**.679**			
Draft item balance	**.669**			
Draft item daily tasks	**.668**	.437		
Draft item fine movement	**.659**			
Draft item get out of bed	**.647**			.472
Draft item need help	**.611**		.485	
Draft item movement	**.608**		.465	
Draft item strength	**.606**		.466	
Draft item concentrate	**.583**		.442	
Draft item open/close fist	**.581**	.448		
Draft item effort	**.564**	.435	.456	
Draft item hobbies	**.562**	.	.471	
Draft item slower	**.549**	.464	.486	
Draft item get up after sitting	**.543**			.490
Draft item severity		**.690**	.401	
Draft item different to usual		**.689**		
Traditional item severity G		**.677**		
Traditional item severity C		**.672**		
Draft item impact		**.628**	.477	
Traditional item duration D		**.626**		
Traditional item severity A	.433	**.625**		
Draft item importance		**.598**	.529	
Draft item timing		**.598**	.529	
Draft item in joints		**.598**	.529	
Draft item coped		-.362		
Draft item worried			**.771**	
Draft item self-conscious			**.743**	
Draft item frustrated	.412		**.679**	
Draft item change plans	.475		**.678**	
Draft item work around	.572		**.602**	
Draft item take longer	.569		**.577**	
Draft item all over				**.764**
Draft item in body				**.672**
Draft item after immobility				**.599**
Draft item sleep	.443			**.451**

Component key: Physical; Severity; Psychosocial; Location

Description of item removal	Rotated PCA output
Rotated PCA 2 (Table 9.6) comprised 37-items and retained the 4-component structure. The Severity and Psychosocial components remained consistent while the Location component was reduced to three items as 'Has RA stiffness affected your sleep?' (Draft item sleep) loaded highest in the Physical component. The lowest loading item was 'Has RA stiffness affected your sleep?' However, internal consistency analyses indicated that the overall Cronbach's alpha coefficient of 0.971 could be improved very slightly (0.972) by the removal of 'Over the past 7 days have you experienced RA stiffness all over?' (Draft item all over) On further consideration, this item had demonstrated the lowest component loading during examination of individual components (Section 9.4.2.2.4). It had also been considered for removal during investigation into item suitability (Chapter 8) where it was identified as having a number (n=17) of poor correlations (<0.3) with other items and a large percentage of responses in one response category (Section 8.5.2.1.3). In contrast, 'Has RA stiffness affected your sleep?' during individual component investigation (Section 9.4.2.2) had demonstrated a slightly higher component loading than 'Over the past 7 days have you experienced RA stiffness all over?' (Table 9.4), and had not been considered for removal previously (Chapter 8). For these reasons, 'Over the past 7 days have you experienced RA stiffness all over?' was removed (Table 9.7).	**Table 9.6: Rotated PCA 2 ('Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' (Draft item coped) removed)**

Item	1	2	3	4
Draft item wash	.792			
Draft item dress	.743			
Draft item grip	.697			
Draft item responsibilities	.684			
Draft item daily tasks	.675	.435		
Draft item balance	.668			
Draft item fine movement	.666			
Draft item get out of bed	.645			.460
Draft item strength	.614		.469	
Draft item need help	.614		.489	
Draft item movement	.613		.471	
Draft item open/close fist	.593	.423		
Draft item concentrate	.583		.445	
Draft item effort	.571	.431	.462	
Draft item hobbies	.564		.477	
Draft item slower	.557	.458	.493	
Draft item get up after sitting	.542			.478
Draft item sleep	.448			.446
Draft item severity		.707	.413	
Traditional item severity G		.695	.401	
Draft item different to usual		.689		
Traditional item severity C	.405	.680		
Traditional item severity A	.440	.646		
Traditional item duration D		.644		
Draft item impact	.406	.643	.488	
Draft item importance		.608	.539	
Draft item timing		.532		.482
Draft item in joints		.530		.465
Draft item worried			.772	
Draft item self-conscious			.746	
Draft item frustrated	.416		.683	
Draft item change plans	.478		.682	
Draft item work around	.575		.606	
Draft item take longer	.574		.583	
Draft item all over				.773
Draft item in body				.669
Draft item after immobility				.579

Component key: Physical; Severity; Psychosocial; Location

Description of item removal	Rotated PCA output
Rotated PCA 3 (Table 9.7) retained a 4-component structure and its Cronbach's alpha coefficient was .972, but the content of the Location component had changed considerably. The two items that remained from the original Location component ('Have you experienced RA stiffness in your body (outside of your joints) over the past 7 days?' (Draft item in body) and 'Over the past 7 days have you experienced RA stiffness after a period of immobility (for example, after sitting for a while)?' (Draft item after immobility)) were joined by four other items ('Has RA stiffness made it difficult to get out of bed?' (Draft item get out of bed), 'Has RA stiffness made it difficult to get up after sitting for a while?' (Draft item get up after sitting), 'Have you had to concentrate to move your body because of RA stiffness?' (Draft item concentrate), and 'Has RA stiffness made it difficult to balance without physically supporting yourself?' (Draft item balance)), which had previously loaded highest on the Physical component. Although this change in content challenged the original label of the Location component, its original name was retained for consistency. As in rotated PCA 2 (Table 9.6), 'Has RA stiffness affected your sleep?' (Draft item sleep) again demonstrated the lowest component loading. This item had also demonstrated inconsistent component placement having loaded (>0.40) on the Location (rotated PCA 1), Physical (rotated PCA 2), and Severity (rotated PCA 3) components. Therefore 'Has RA stiffness affected your sleep?' was removed (Table 9.8).	**Table 9.7: Rotated PCA 3 ('Over the past 7 days have you experienced RA stiffness all over?' (Draft item all over) removed)**

Item	1	2	3	4
Draft item grip	**.718**			
Draft item open/close fist	**.703**			
Draft item fine movement	**.685**			
Draft item wash	**.661**			.492
Draft item responsibilities	**.641**			
Draft item strength	**.623**		.470	
Draft item daily tasks	**.621**	.431		
Draft item dress	**.593**			.539
Draft item need help	**.551**		.501	
Draft item slower	**.531**	.436	.510	
Draft item movement	**.515**		.497	
Draft item effort	**.494**	.434	.484	
Draft item severity		**.708**	.448	
Traditional item severity G		**.689**	.439	
Traditional item severity C		**.687**		
Draft item different to usual		**.669**		
Traditional item severity A		**.651**		.407
Draft item impact		**.647**	.519	
Draft item timing		**.644**		
Traditional item duration D		**.628**		
Draft item in joints		**.625**		
Draft item importance		**.613**	.569	
Draft item sleep		**.434**		.429
Draft item worried			**.777**	
Draft item self-conscious			**.755**	
Draft item frustrated			**.701**	
Draft item change plans	.458		**.687**	
Draft item work around	.538		**.612**	
Draft item take longer	.519		**.602**	
Draft item hobbies	.466		**.499**	
Draft item get up after sitting				**.692**
Draft item get out of bed	.406			**.679**
Draft item after immobility		.473		**.598**
Draft item concentrate			.491	**.496**
Draft item balance	.484			**.491**
Draft item in body		.460		**.471**

Component key: Physical; Severity; Psychosocial; Location

Description of item removal	Rotated PCA output
Rotated PCA 4 (Table 9.8) retained a 4-component structure and had a Cronbach's alpha coefficient of 0.971. The structure was similar to the solution identified in rotated PCA 3 although the Physical and Severity components changed places. The content of the Physical and Psychosocial components remained consistent while the Severity component gained 'Have you experienced RA stiffness in your body (outside of your joints) over the past 7 days?' (Draft item in body) which originally loaded highest in the Location component. This item also demonstrated the lowest overall component loading. As a result, 'Have you experienced RA stiffness in your body (outside of your joints) over the past 7 days?' was removed (Table 9.9).	**Table 9.8: Rotated PCA 4 ('Has RA stiffness affected your sleep?' (Draft item sleep) removed)**

Item	1	2	3	4
Draft item severity	**.715**		.445	
Traditional item severity G	**.697**		.435	
Traditional item severity C	**.694**			
Traditional item severity A	**.663**			
Draft item different to usual	**.661**			
Draft item timing	**.657**			
Draft item impact	**.652**		.517	
Draft item in joints	**.637**			
Traditional item duration D	**.629**			
Draft item importance	**.619**		.568	
Draft item in body	**.482**			.437
Draft item grip		**.717**		
Draft item open/close fist		**.710**		
Draft item fine movement		**.684**		
Draft item wash		**.643**		.514
Draft item responsibilities		**.634**		
Draft item strength		**.621**	.469	
Draft item daily tasks	.436	**.614**		
Draft item dress		**.576**		.553
Draft item need help		**.543**	.500	
Draft item slower	.436	**.530**	.509	
Draft item movement		**.504**	.491	
Draft item effort	.440	**.488**	.484	
Draft item worried			**.780**	
Draft item self-conscious			**.756**	
Draft item frustrated			**.702**	
Draft item change plans		.454	**.690**	
Draft item work around		.533	**.613**	
Draft item take longer		.514	**.600**	
Draft item hobbies		.457	**.498**	
Draft item get up after sitting				**.697**
Draft item get out of bed				**.681**
Draft item after immobility	.502			**.579**
Draft item balance		.461		**.530**
Draft item concentrate			.480	**.522**

Component key: Physical; Severity; Psychosocial; Location

Description of item removal	Rotated PCA output
Rotated PCA 5 (Table 9.9) retained a 4-component structure and its Cronbach's alpha coefficient was 0.971. Again the structure was similar to the previous two solutions (rotated PCA 3 and 4), however the Physical, Psychosocial and Severity components changed places. The Severity component remained consistent in content while the other components varied slightly as a result of changes in the highest loading of some items. The Psychosocial component gained 'Has your movement been restricted because of RA stiffness?' (Draft item movement) and 'Have your daily tasks and activities required more effort because of RA stiffness?' (Draft item effort) from the Physical component. The Physical component gained 'Has RA stiffness made it difficult to get out of bed?' (Draft item get out of bed) from the Location component but lost 'Has RA stiffness made it difficult to dress or undress yourself?' (Draft item dress) and 'Has RA stiffness made it difficult to wash yourself (for example, have a shower)?' (Draft item dress) back to the Location component. Overall, 'Have your daily tasks and activities required more effort because of RA stiffness?' was the lowest loading item and was removed (Table 9.10).	**Table 9.9: Rotated PCA 5 ('Have you experienced RA stiffness in your body (outside of your joints) over the past 7 days?' (Draft item in body) removed)**

Item	1	2	3	4
Draft item severity	**.723**	.448		
Traditional item severity G	**.711**	.438		
Traditional item severity C	**.708**			
Traditional item severity A	**.682**			.425
Draft item timing	**.665**			
Draft item impact	**.659**	.520		
Draft item different to usual	**.655**			
Draft item in joints	**.647**			
Traditional item duration D	**.629**			
Draft item importance	**.627**	.570		
Draft item worried		**.779**		
Draft item self-conscious		**.753**		
Draft item frustrated		**.700**		
Draft item change plans		**.687**	.446	
Draft item work around		**.608**	.518	
Draft item take longer		**.597**	.487	
Draft item hobbies		**.496**	.429	.426
Draft item movement		**.488**	.467	.436
Draft item effort	.442	**.482**	.467	
Draft item grip			**.716**	
Draft item open/close fist			**.704**	
Draft item fine movement			**.680**	
Draft item strength		.463	**.624**	
Draft item responsibilities			**.601**	
Draft item daily tasks	.437		**.586**	.408
Draft item need help		.495	**.518**	
Draft item slower	.436	.507	**.512**	
Draft item get out of bed				**.716**
Draft item get up after sitting				**.710**
Draft item dress			.506	**.610**
Draft item balance				**.604**
Draft item wash			.558	**.598**
Draft item concentrate		.480		**.572**
Draft item after immobility	.523			**.540**

Component key: Physical; Severity; Psychosocial; Location

Description of item removal	Rotated PCA output
Rotated PCA 6 (Table 9.10) contained 33-items across 4-components, with an overall Cronbach's alpha coefficient of 0.969. This solution was comparable to the rotated solution produced in rotated PCA 5 except that in the Physical and Location components, 'Has RA stiffness made it difficult to get out of bed?' (Draft item get out of bed) now loaded highest in the Location component. 'Has your movement been restricted because of RA stiffness?' (Draft item movement) had the lowest component loading overall, therefore it was removed (Table 9.11).	**Table 9.10: Rotated PCA 6 ('Have your daily tasks and activities required more effort because of RA stiffness?' (Draft item effort) removed)**

Item	1	2	3	4
Draft item severity	**.724**	.447		
Traditional item severity G	**.712**	.438		
Traditional item severity C	**.709**			
Traditional item severity A	**.683**			.426
Draft item timing	**.665**			
Draft item impact	**.660**	.518		
Draft item different to usual	**.656**			
Draft item in joints	**.647**			
Traditional item duration D	**.630**			
Draft item importance	**.628**	.569		
Draft item worried		**.779**		
Draft item self-conscious		**.755**		
Draft item frustrated		**.700**		
Draft item change plans		**.686**	.444	
Draft item work around		**.608**	.517	
Draft item take longer		**.594**	.482	
Draft item hobbies		**.494**	.424	.429
Draft item movement		**.488**	.465	.438
Draft item grip			**.715**	
Draft item open/close fist			**.705**	
Draft item fine movement			**.681**	
Draft item strength		.464	**.624**	
Draft item responsibilities			**.597**	.400
Draft item daily tasks	.438		**.581**	.411
Draft item need help		.497	**.520**	
Draft item slower	.436	.504	**.508**	
Draft item get out of bed				**.717**
Draft item get up after sitting				**.711**
Draft item dress			.505	**.611**
Draft item balance				**.605**
Draft item wash			.557	**.600**
Draft item concentrate		.479		**.573**
Draft item after immobility	.523			**.540**

Component key: Physical; Severity; Psychosocial; Location

Description of item removal	Rotated PCA output
Rotated PCA 7 (Table 9.11) retained a 4-component structure and had an overall Cronbach's alpha coefficient of 0.968. Again the Severity and Psychosocial components remained consistent however, the Physical component lost 'Has RA stiffness made you slower (for example, unable to do things quickly)?' (Draft item slower) to the Location component. 'Has RA stiffness made it difficult to do hobbies or activities you enjoy?' (Draft item hobbies) had the lowest component loading overall and was removed (Table 9.12).	**Table 9.11: Rotated PCA 7 ('Has your movement been restricted because of RA stiffness?' (Draft item movement) removed)**

Item	1	2	3	4
Draft item severity	**.724**	.446		
Traditional item severity G	**.713**	.437		
Traditional item severity C	**.708**			
Traditional item severity A	**.682**			.428
Draft item timing	**.666**			
Draft item impact	**.660**	.519		
Draft item different to usual	**.657**			
Draft item in joints	**.646**			
Traditional item duration D	**.629**			
Draft item importance	**.628**	.570		
Draft item worried		**.780**		
Draft item self-conscious		**.755**		
Draft item frustrated		**.700**		
Draft item change plans		**.687**	.443	
Draft item work around		**.609**	.517	
Draft item take longer		**.594**	.481	
Draft item hobbies		**.493**	.422	.429
Draft item grip			**.715**	
Draft item open/close fist			**.704**	
Draft item fine movement			**.681**	
Draft item strength		.462	**.622**	
Draft item responsibilities			**.597**	.404
Draft item daily tasks	.437		**.581**	.414
Draft item need help		.498	**.519**	
Draft item slower	.437	.504	**.507**	
Draft item get out of bed				**.721**
Draft item get up after sitting				**.709**
Draft item dress			.504	**.613**
Draft item balance				**.604**
Draft item wash			.556	**.603**
Draft item concentrate		.479		**.571**
Draft item after immobility	.527			**.535**

Component key: Physical; Severity; Psychosocial; Location

Description of item removal	Rotated PCA output

Rotated PCA 8 (Table 9.12) contained 31-items and its overall Cronbach's alpha coefficient was 0.966. However, the 4-component solution reduced to a 3-component solution where the original Location component was lost. The content of the three remaining components incorporated items that had loaded highest in the Location component in the previous solution. All items apart from 'Has RA stiffness made it difficult to open and close your fist?' (Draft item open/close fist) had component loadings defined as 'good' or better which was an improvement on previous solutions. The lowest loading item was 'Has RA stiffness made it difficult to open and close your fist?' However on consideration of internal consistency analyses, Cronbach's alpha coefficient (0.966) could have been increased very slightly (0.967) by removing 'Over the past 7 days has your RA stiffness been different to usual for you?' (Draft item different to usual). This item was the second lowest loading during examination of individual components (Section 9.4.2.2.2). It had also been considered for removal during investigation into item suitability (Chapter 8) where it was identified as having a number (n=11) of poor correlations (<0.3) with other items and a large percentage of responses in one response category (Section 8.5.2.1.3). 'Has RA stiffness made it difficult to open and close your fist?' had the lowest component loading and Cronbach alpha coefficient during examination of individual components. It also was the only item to demonstrate any weak correlations with other items within the component (Section 9.4.2.2.1). Both items had evidence to justify their removal. However, as 'Over the past 7 days has your RA stiffness been different to usual for you?' had evidence of poor suitability (Chapter 8) it was removed first (Table 9.13).

Table 9.12: Rotated PCA 8 ('Has RA stiffness made it difficult to do hobbies or activities you enjoy?' (Draft item hobbies) removed)

Item	1	2	3
Draft item wash	.793		
Draft item dress	.756		
Draft item get out of bed	.693		
Draft item grip	.662		.408
Draft item balance	.662		
Draft item responsibilities	.662		.407
Draft item daily tasks	.653	.467	
Draft item fine movement	.638		
Draft item need help	.594		.544
Draft item get up after sitting	.588	.473	
Draft item concentrate	.580		.459
Draft item strength	.574		.541
Draft item open/close fist	.517		
Draft item severity		.750	.462
Traditional item severity G		.743	.447
Traditional item severity C		.735	
Traditional item severity A	.426	.725	
Draft item impact		.686	.537
Draft item timing		.685	
Draft item in joints		.670	
Draft item importance		.654	.581
Traditional item duration D		.637	
Draft item different to usual		.635	
Draft item after immobility		.594	
Draft item worried			.780
Draft item self-conscious			.752
Draft item change plans	.438		.725
Draft item frustrated			.705
Draft item work around	.545		.655
Draft item take longer	.535		.631
Draft item slower	.512	.454	.556

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
The rotated PCA 9 (Table 9.13) contained 30-items and retained a 3-component solution with the Cronbach's alpha coefficient also remaining stable at 0.967. The components remained steady with rotated PCA 8 (Table 9.12) apart from 'Has RA stiffness made it difficult to get up after sitting for a while?' (Draft item get up after sitting) which loaded highest within the Severity component. Consistent with rotated PCA 8 (Table 9.12), 'Has RA stiffness made it difficult to open and close your fist?' (Draft item open/close fist) demonstrated the lowest component loading. Given the poor performance of 'Has RA stiffness made it difficult to open and close your fist?' discussed above, it was removed (Table 9.14).	Table 9.13: Rotated PCA 9 ('Over the past 7 days has your RA stiffness been different to usual for you?' (Draft item different to usual) removed)

Item	1	2	3
Draft item wash	**.789**		
Draft item dress	**.745**	.424	
Draft item grip	**.672**		
Draft item responsibilities	**.669**		
Draft item balance	**.660**		
Draft item daily tasks	**.659**	.472	
Draft item get out of bed	**.657**	.440	
Draft item fine movement	**.654**		
Draft item need help	**.598**		.526
Draft item strength	**.585**		.523
Draft item concentrate	**.570**		.449
Draft item open/close fist	**.538**		
Traditional item severity A		**.760**	
Draft item severity		**.752**	.463
Traditional item severity G		**.752**	.446
Traditional item severity C		**.750**	
Draft item timing		**.716**	
Draft item in joints		**.701**	
Draft item impact		**.697**	.534
Draft item importance		**.667**	.579
Draft item after immobility		**.638**	
Traditional item duration D		**.637**	
Draft item get up after sitting	.537	**.539**	
Draft item worried			**.780**
Draft item self-conscious			**.747**
Draft item change plans	.456		**.715**
Draft item frustrated		.403	**.695**
Draft item work around	.554		**.638**
Draft item take longer	.535	.407	**.617**
Draft item slower	.513	.468	**.543**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 10 (Table 9.14) contained 3-components and had an overall Cronbach's alpha coefficient of 0.967. Consistent with rotated PCA 9 (Table 9.13), all item loadings could be defined as 'good' or better. There were slight changes to the solution including 'Has RA stiffness made it difficult to get up after sitting for a while?' (Draft item get up after sitting) loading highest in Physical component instead of the Severity component. These components also swapped places. 'Has RA stiffness reduced your strength to do tasks?' (Draft item strength) loaded highest within the Psychosocial component rather than the Physical component. Overall, 'Has RA stiffness reduced your strength to do tasks?' was the lowest loading item and was removed (Table 9.15).	**Table 9.14: Rotated PCA 10 ('Has RA stiffness made it difficult to open and close your fist?' (Draft item open/close fist) removed)**

Item	1	2	3
Traditional item severity A	**.757**		
Draft item severity	**.754**		.468
Traditional item severity G	**.752**		.451
Traditional item severity C	**.751**		
Draft item timing	**.718**		
Draft item in joints	**.701**		
Draft item impact	**.699**		.541
Draft item importance	**.669**		.583
Traditional item duration D	**.644**		
Draft item after immobility	**.628**		
Draft item wash		**.789**	
Draft item dress	.415	**.749**	
Draft item get out of bed	.425	**.684**	
Draft item balance		**.676**	
Draft item responsibilities		**.646**	.421
Draft item daily tasks	.470	**.641**	
Draft item grip		**.624**	.429
Draft item fine movement		**.610**	
Draft item concentrate		**.580**	.454
Draft item need help		**.580**	.551
Draft item get up after sitting	.524	**.568**	
Draft item worried			**.779**
Draft item self-conscious			**.747**
Draft item change plans		.432	**.732**
Draft item frustrated	.405		**.702**
Draft item work around		.527	**.662**
Draft item take longer	.408	.512	**.637**
Draft item slower	.472	.483	**.566**
Draft item strength		.536	**.558**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 11 (Table 9.15) comprised 28-items over 3-components and had an overall Cronbach's alpha coefficient of 0.965. All components remained consistent and 'Has RA stiffness made you slower (for example, unable to do things quickly)?' (Draft item slower) was the lowest loading item. Both 'Has RA stiffness made you slower (for example, unable to do things quickly)?' and 'Has it taken you longer to do your daily tasks or activities because of RA stiffness?' (Draft item take longer) also demonstrated acceptable component loadings (>0.45) on all components which can cause difficulties with interpretation. Therefore, 'Has RA stiffness made you slower (for example, unable to do things quickly)?' was removed (Table 9.16).	**Table 9.15: Rotated PCA 11 ('Has RA stiffness reduced your strength to do tasks?' (Draft item strength) removed)**

Item	1	2	3
Traditional item severity A	**.751**		
Draft item severity	**.751**		.469
Traditional item severity C	**.750**		
Traditional item severity G	**.748**		.452
Draft item timing	**.723**		
Draft item in joints	**.710**		
Draft item impact	**.697**		.540
Draft item importance	**.665**		.584
Traditional item duration D	**.643**		
Draft item after immobility	**.627**		
Draft item wash		**.796**	
Draft item dress	.409	**.755**	
Draft item get out of bed	.417	**.689**	
Draft item balance		**.682**	
Draft item responsibilities		**.652**	.419
Draft item daily tasks	.468	**.646**	
Draft item grip		**.620**	.406
Draft item fine movement		**.610**	
Draft item concentrate		**.586**	.454
Draft item need help		**.581**	.539
Draft item get up after sitting	.519	**.571**	
Draft item worried			**.783**
Draft item self-conscious			**.749**
Draft item change plans		.443	**.733**
Draft item frustrated	.401		**.702**
Draft item work around		.532	**.654**
Draft item take longer	.409	.518	**.631**
Draft item slower	.477	.485	**.556**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 12 (Table 9.16) retained 3-components and had a Cronbach's alpha coefficient of 0.964. The components again remained consistent. Although 'Has it taken you longer to do your daily tasks or activities because of RA stiffness?' (Draft item take longer) continued to demonstrate acceptable component loadings (>0.45) on all components, 'Has RA stiffness made it difficult to get up after sitting for a while?' (Draft item get up after sitting) had the lowest component loading overall and was removed (Table 9.17).	**Table 9.16: Rotated PCA 12 ('Has RA stiffness made you slower (for example, unable to do things quickly)?' (Draft item slower) removed)**

Item	1	2	3
Draft item severity	**.753**		.461
Traditional item severity A	**.753**		
Traditional item severity C	**.752**		
Traditional item severity G	**.750**		.446
Draft item timing	**.724**		
Draft item in joints	**.711**		
Draft item impact	**.700**		.532
Draft item importance	**.668**		.577
Traditional item duration D	**.646**		
Draft item after immobility	**.627**		
Draft item wash		**.798**	
Draft item dress	.410	**.756**	
Draft item get out of bed	.418	**.688**	
Draft item balance		**.685**	
Draft item responsibilities		**.659**	.406
Draft item daily tasks	.471	**.653**	
Draft item grip		**.629**	
Draft item fine movement		**.617**	
Draft item need help		**.589**	.528
Draft item concentrate		**.589**	.449
Draft item get up after sitting	.520	**.569**	
Draft item worried			**.786**
Draft item self-conscious			**.754**
Draft item change plans		.454	**.724**
Draft item frustrated	.404		**.702**
Draft item work around		.541	**.644**
Draft item take longer	.413	.527	**.615**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 13 (Table 9.17) contained 26-items and maintained the 3-component structure. The overall Cronbach's alpha coefficient was 0.963. The components remained consistent. 'Have you had to concentrate to move your body because of RA stiffness?' (Draft item concentrate) had the lowest loading and was removed (Table 9.18).	**Table 9.17: Rotated PCA 13 ('Has RA stiffness made it difficult to get up after sitting for a while?' (Draft item get up after sitting) removed)**

Item	1	2	3
Draft item severity	**.759**		.444
Traditional item severity A	**.757**		
Traditional item severity C	**.756**	.401	
Traditional item severity G	**.755**		.433
Draft item timing	**.726**		
Draft item in joints	**.714**		
Draft item impact	**.705**		.519
Draft item importance	**.673**		.564
Traditional item duration D	**.651**		
Draft item after immobility	**.619**		
Draft item wash		**.808**	
Draft item dress	.415	**.759**	
Draft item responsibilities	.404	**.682**	
Draft item daily tasks	.480	**.676**	
Draft item balance		**.668**	
Draft item get out of bed	.416	**.665**	
Draft item grip		**.656**	
Draft item fine movement		**.640**	
Draft item need help		**.604**	.509
Draft item concentrate		**.565**	.472
Draft item worried			**.789**
Draft item self-conscious			**.766**
Draft item frustrated	.405		**.707**
Draft item change plans		.474	**.704**
Draft item work around		.557	**.625**
Draft item take longer	.417	.537	**.602**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 14 (Table 9.18) retained the 3-component solution and had a Cronbach's alpha coefficient of 0.962. All components continued to remain consistent. Overall, 'Has it taken you longer to do your daily tasks or activities because of RA stiffness?' (Draft item take longer) had the lowest loading, and had consistently demonstrated loading >0.45 on all components in previous solutions, therefore it was removed (Table 9.19).	**Table 9.18: Rotated PCA 14 ('Have you had to concentrate to move your body because of RA stiffness?' (Draft item concentrate) removed)**

Item	1	2	3
Draft item severity	**.762**		.441
Traditional item severity G	**.757**		.431
Traditional item severity A	**.754**	.403	
Traditional item severity C	**.754**		
Draft item timing	**.721**		
Draft item in joints	**.709**		
Draft item impact	**.705**		.518
Draft item importance	**.673**		.564
Traditional item duration D	**.657**		
Draft item after immobility	**.625**		
Draft item wash		**.809**	
Draft item dress	.409	**.763**	
Draft item responsibilities		**.693**	
Draft item daily tasks	.469	**.686**	
Draft item grip		**.664**	
Draft item get out of bed	.424	**.654**	
Draft item fine movement		**.648**	
Draft item balance		**.639**	
Draft item need help		**.607**	.515
Draft item worried			**.790**
Draft item self-conscious			**.767**
Draft item frustrated		.478	**.710**
Draft item change plans	.406		**.708**
Draft item work around		.563	**.631**
Draft item take longer	.416	.537	**.604**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 15 (Table 9.19) retained 3-components and had a Cronbach's alpha coefficient of 0.959. 'Have you needed help (from others or gadgets) because of RA stiffness?' (Draft item need help) and 'Have you had to work around your RA stiffness (or do things in a different way)?' (Draft item work around) loaded lowest and both had a loading of 0.615 on their respective components. To identify which of these items to remove, internal consistency analyses were inspected. The removal of either item would reduce Cronbach's alpha coefficient slightly to .958 thus not differentiating between items. However, 'Have you had to work around your RA stiffness (or do things in a different way)?' had a slightly higher corrected item total correlation (*r*=0.787) than 'Have you needed help (from others or gadgets) because of RA stiffness?' (*r*=0.760), therefore 'Have you needed help (from others or gadgets) because of RA stiffness?' was removed (Table 9.20).	**Table 9.19: Rotated PCA 15 ('Has it taken you longer to do your daily tasks or activities because of RA stiffness?' (Draft item take longer) 10.30 removed)**

Item	1	2	3
Draft item severity	**.766**		.430
Traditional item severity G	**.761**		.420
Traditional item severity C	**.755**		
Traditional item severity A	**.755**	.408	
Draft item timing	**.720**		
Draft item in joints	**.711**		
Draft item impact	**.710**		.505
Draft item importance	**.677**		.555
Traditional item duration D	**.661**		
Draft item after immobility	**.627**		
Draft item wash		**.812**	
Draft item dress	.408	**.766**	
Draft item responsibilities		**.698**	
Draft item daily tasks	.474	**.690**	
Draft item grip		**.669**	
Draft item get out of bed	.419	**.657**	
Draft item fine movement		**.653**	
Draft item balance		**.644**	
Draft item need help		**.615**	.502
Draft item worried			**.797**
Draft item self-conscious			**.771**
Draft item frustrated	.410		**.701**
Draft item change plans		.488	**.696**
Draft item work around		.571	**.615**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 16 (Table 9.20) contained 23-items in a 3-component solution which had a Cronbach's alpha coefficient of 0.958. All components continued to remain consistent. As identified in rotated PCA 15 (Table 9.19), 'Have you had to work around your RA stiffness (or do things in a different way)?' (Draft item work around) still had the lowest component loading, therefore it was removed (Table 9.21).	**Table 9.20: Rotated PCA 16 ('Have you needed help (from others or gadgets) because of RA stiffness?' (Draft item need help) removed)**

Item	1	2	3
Draft item severity	**.765**		.432
Traditional item severity G	**.761**		.421
Traditional item severity A	**.750**	.411	
Traditional item severity C	**.749**		
Draft item timing	**.725**		
Draft item in joints	**.710**		
Draft item impact	**.709**		.507
Draft item importance	**.676**		.557
Traditional item duration D	**.652**		
Draft item after immobility	**.632**		
Draft item wash		**.819**	
Draft item dress		**.770**	
Draft item responsibilities		**.708**	
Draft item daily tasks	.469	**.692**	
Draft item get out of bed	.407	**.666**	
Draft item grip		**.657**	
Draft item fine movement		**.648**	
Draft item balance		**.640**	
Draft item worried			**.806**
Draft item self-conscious			**.771**
Draft item frustrated	.408		**.705**
Draft item change plans		.485	**.701**
Draft item work around		.568	**.620**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 17 (Table 9.21) contained 22-items in a 3-component solution which had a Cronbach's alpha coefficient of 0.956. The components retained their structure and it was identified that item 'Over the past 7 days have you experienced RA stiffness after a period of immobility (for example, after sitting for a while)?' (Draft item after immobility) had the lowest component loading, therefore it was removed (Table 9.22).	**Table 9.21: Rotated PCA 17 ('Have you had to work around your RA stiffness (or do things in a different way)?' (Draft item work around) removed)**

Item	1	2	3
Draft item severity	**.754**		.443
Traditional item severity G	**.752**		.431
Traditional item severity C	**.743**		
Traditional item severity A	**.739**	.419	
Draft item timing	**.738**		
Draft item in joints	**.720**		
Draft item impact	**.699**		.514
Draft item importance	**.662**		.568
Traditional item duration D	**.635**		
Draft item after immobility	**.632**		
Draft item wash		**.826**	
Draft item dress		**.777**	
Draft item responsibilities		**.715**	
Draft item daily tasks	.468	**.696**	
Draft item get out of bed		**.674**	
Draft item grip		**.661**	
Draft item fine movement		**.653**	
Draft item balance		**.652**	
Draft item worried			**.813**
Draft item self-conscious			**.774**
Draft item frustrated			**.702**
Draft item change plans		.492	**.673**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 18 (Table 9.22) retained the 3-component solution with a Cronbach's alpha coefficient of 0.955. In addition to consistency of the structure of the solution, all item loadings could be defined as 'very good' or 'excellent'. Despite this, for the purpose of testing item removal, the process was continued and 'Has RA stiffness made it difficult to balance without physically supporting yourself?' (Draft item balance) was identified as having the lowest component loading, and was removed (Table 9.23).	**Table 9.22: Rotated PCA 18 ('Over the past 7 days have you experienced RA stiffness after a period of immobility (for example, after sitting for a while)?' (Draft item after immobility) removed)**

Item	1	2	3
Draft item severity	**.770**		.427
Traditional item severity G	**.763**		.417
Traditional item severity C	**.756**		
Traditional item severity A	**.746**	.423	
Draft item timing	**.716**		
Draft item in joints	**.710**		
Draft item impact	**.710**		.501
Draft item importance	**.675**		.556
Traditional item duration D	**.661**		
Draft item wash		**.827**	
Draft item dress		**.781**	
Draft item responsibilities		**.713**	
Draft item daily tasks	.472	**.697**	
Draft item get out of bed		**.682**	
Draft item grip		**.662**	
Draft item fine movement		**.653**	
Draft item balance		**.650**	
Draft item worried			**.812**
Draft item self-conscious			**.786**
Draft item frustrated			**.705**
Draft item change plans		.487	**.672**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 19 (Table 9.23) contained 20-items and retained the 3-component solution with a Cronbach's alpha coefficient of 0.954. Component consistency was retained and 'Has RA stiffness made it difficult to do fine movements (for example, write with a pen)?' (Draft item fine movement) had the lowest component loading, and was removed (Table 9.24).	**Table 9.23: Rotated PCA 19 ('Has RA stiffness made it difficult to balance without physically supporting yourself?' (Draft item balance) removed)**

Item	1	2	3
Draft item severity	.779		.428
Traditional item severity G	.773		.418
Traditional item severity C	.756		
Traditional item severity A	.747	.419	
Draft item impact	.715		.505
Draft item in joints	.702		
Draft item timing	.695		
Draft item importance	.676		.559
Traditional item duration D	.671		
Draft item wash		.818	
Draft item dress		.786	
Draft item responsibilities		.714	
Draft item daily tasks	.463	.700	
Draft item grip		.673	
Draft item get out of bed		.672	
Draft item fine movement		.668	
Draft item worried			.816
Draft item self-conscious			.792
Draft item frustrated			.711
Draft item change plans		.477	.683

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Rotated PCA 20 (Table 9.24) contained 19-items and retained the consistent 3-component solution with a Cronbach's alpha coefficient of 0.952. 'Has RA stiffness made it difficult to grip or hold things?' (Draft item grip) had the lowest loading, thus it was removed (Table 9.25).	**Table 9.24: Rotated PCA 20 ('Has RA stiffness made it difficult to do fine movements (for example, write with a pen)?' (Draft item fine movement) removed)**

Item	1	2	3
Draft item severity	**.778**		.434
Traditional item severity G	**.773**		.424
Traditional item severity C	**.752**		
Traditional item severity A	**.741**	.426	
Draft item impact	**.716**		.511
Draft item in joints	**.710**		
Draft item timing	**.705**		
Draft item importance	**.676**		.565
Traditional item duration D	**.672**		
Draft item wash		**.830**	
Draft item dress		**.794**	
Draft item responsibilities		**.725**	
Draft item daily tasks	.464	**.698**	
Draft item get out of bed		**.687**	
Draft item grip		**.620**	
Draft item worried			**.819**
Draft item self-conscious			**.800**
Draft item frustrated			**.718**
Draft item change plans		.458	**.694**

Component key: Physical; Severity; Psychosocial

Description of item removal	Rotated PCA output
Following the removal of item 10.21, the 3-component solution that had been stable for 13 rounds of item removal, reduced to a 2-component solution containing 18-items. Although the internal consistency remained high with a Cronbach's alpha coefficient of 0.951, 12 items loaded on both the Physical and Severity components, which made interpretation difficult. Here it appeared that the Severity component remained stable from the solution generated in rotated PCA 20 (Table 9.24). However, the Physical and Psychosocial components appeared to merge into one component. This was labelled as the Physical component given the slightly higher number of items from the earlier defined Physical component. Testing of item removal was continued with the removal of the lowest loading item in each round. The 2-component structure was sustained for six further rounds of item removal after which a single component solution was formed containing 11-items. Subsequent PCA solutions have not been presented as given the above solutions it appeared likely that a multi-dimensional solution was present and would be most appropriate to represent the available items.	**Table 9.25: Rotated PCA 21 ('Has RA stiffness made it difficult to grip or hold things?' (Draft item grip) removed)**

Item	1	2
Draft item severity	**.799**	.495
Traditional item severity C	**.796**	.419
Traditional item severity G	**.796**	.488
Traditional item severity A	**.789**	.488
Draft item in joints	**.767**	
Draft item impact	**.739**	.576
Draft item timing	**.736**	
Draft item importance	**.690**	.605
Traditional item duration D	**.670**	
Draft item change plans		**.811**
Draft item self-conscious		**.811**
Draft item worried		**.795**
Draft item frustrated	.417	**.773**
Draft item responsibilities	.500	**.699**
Draft item wash	.433	**.676**
Draft item dress	.516	**.651**
Draft item daily tasks	.586	**.635**
Draft item get out of bed	.488	**.605**

Component key: Physical; Severity

9.4.2.3.1 Review item removal

In the initial 4-component structure, components 1, 2 and 3 ("Physical", "Severity", and "Psychosocial") not only had a strong statistical basis but also appeared to generally capture items assessing related concepts (Section 9.4.2.2). However, component 4 ("Location") had weaker statistical evidence and conceptual clarity. During early rounds of item removal, three items (7.4, 8.8, and 9.9) originally contained within "Location" were removed indicating that this component contained poorly performing items and the content was inconsistent. The revised content of the component (Tables 9.7-9.11) at that stage did not lead to clarity from a conceptual perspective as items that loaded highest within the component were inconsistent across iterations and it did not appear to capture a unified concept. After this, a conceptually coherent 3-component structure emerged (Table 9.12) and remained for 13 rounds of item removal. The component loadings could be described as 'good' or better, improving in later solutions and there was less movement of items between components than seen during the first seven rounds of item removal.

The 2-component solutions generated after PCA 21 (Table 9.25) retained the conceptual essence of the 3-component solution but did not make it clearer. The merging of the "Physical" and "Psychosocial" components and the large number of items that loaded substantively on both components, reduced distinctiveness and interpretation from a conceptual perspective. Although item reduction was pursued further to test whether a smaller and clearer item structure resulted, this did not happen. Given the consistency and conceptual relevance of the 3-component solution, its individual components were re-examined.

9.4.2.3.2 Re-examine each component individually (3-component solution)

The first iteration of the 3-component solution (Table 9.12) was used to re-examine each component.

9.4.2.4.1 Component 1 ("Physical")

A single component was identified and all item loadings were 'excellent' (Table 9.26). The Cronbach's alpha coefficient for the 13-items was .960 which could not be improved by item removal. This was only slightly lower than that for "Physical" in the 4-component solution (0.973). Correlations between items were moderate or strong. However, correlations between 'Has RA stiffness made it difficult to open and close your fist?' and 'Has RA stiffness made it difficult to get out of bed?', 'Has RA stiffness

made it difficult to get up after sitting for a while?' and 'Has RA stiffness made it difficult to balance without physically supporting yourself?' were 'weak' (r=≤0.46). 'Has RA stiffness made it difficult to open and close your fist?' had the lowest loading and Cronbach's alpha coefficient.

Table 9.26: Item loadings and Cronbach's alpha coefficient for component 1 ("Physical")

Item	Loading	Cronbach's alpha coefficient
Draft item daily tasks	.876	.850
Draft item dress	.868	.838
Draft item responsibilities	.866	.827
Draft item wash	.854	.828
Draft item grip	.847	.831
Draft item strength	.836	.821
Draft item need help	.820	.795
Draft item get out of bed	.808	.766
Draft item fine movement	.807	.773
Draft item concentrate	.790	.760
Draft item get up after sitting	.780	.750
Draft item balance	.776	.744
Draft item open/close fist	.720	.682

9.4.2.4.2 Component 2 ("Severity")

A single component was identified and all item loadings were 'excellent' apart from item 8.5 which was 'good' (Table 9.27). The Cronbach's alpha coefficient for the 11-items was 0.935 which was slightly higher than that for "Physical" in the 4-component solution (0.926). The separate removal of each of three items would slightly improve the Cronbach's alpha coefficient (7.3=0.938, 8.5=0.938, and 8.7=0.936). Correlations between items were generally moderate or strong. However, correlations were weak between 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?' and 'Over the past 7 days have you experienced RA stiffness after a period of immobility (for example, after sitting for a while)?' (r=0.459) and 'Over the past 7 days when have you experienced RA stiffness?' (r=0.479). 'Over the past 7 days has your RA stiffness been different to usual for you?' demonstrated weak correlations with eight of the items (r=≤0.497).

Table 9.27: Item loadings and Cronbach's alpha coefficient for component 2 ("Severity")

Item	Loading	Cronbach's alpha coefficient
Draft item severity	.951	.949
Traditional item severity G	.946	.939
Draft item impact	.941	.935
Traditional item severity A	.925	.909
Draft item importance	.918	.911
Traditional item severity C	.895	.865
Traditional item duration D	.764	.930
Draft item in joints	.734	.636
Draft item after immobility	.724	.654
Draft item timing	.710	.621
Draft item different to usual	.582	.536

9.4.2.4.3 Component 3 ("Psychosocial")

A single component was identified and all item loadings were 'excellent' (Table 9.28). The Cronbach's alpha coefficient for the 7-item component was 0.948 which could not be improved by item removal. This was slightly higher than that for "Psychosocial" in the 4-component solution (0.941). All correlations between items were moderate or strong.

Table 9.28: Item loadings and Cronbach's alpha coefficient for component 3 ("Psychosocial")

Item	Loading	Cronbach's alpha coefficient
Draft item take longer	.913	.936
Draft item work around	.902	.937
Draft item frustrated	.896	.937
Draft item change plans	.884	.939
Draft item slower	.864	.941
Draft item self-conscious	.834	.944
Draft item worried	.819	.946

9.4.2.4 Decide on the final component structure

The first 3-component solution (Table 9.12) was identified as being consistent and effective when considering both the whole structure and individual components. Subsequent testing of the 3-component solution indicated that it was statistically stable and conceptually sound. It was next important to review the above solutions in order to decide which of the 3-component solutions should be retained to ensure that the chosen solution was as good as it could be both statistically and conceptually. In all 3-component solutions there were a number of items within "Severity" that

appeared to be capturing similar information. Two items both captured MS severity asking 'How would you describe the overall level of morning stiffness you have had from the time you wake up?' but were measured on different scales (NRS and 5-option ordinal scale) while 'Please circle the number that best describes the severity of your RA stiffness over the past 7 days' and 'Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week' both captured stiffness severity on NRS but were worded differently. Given this, it was important to test the removal of conceptually similar items to identify the smallest component structure. Secondly, the components contained different numbers of items yet did not differ in importance. Therefore, the removal of items was influenced by consideration of the number of items per component.

To consider these aspects, the 3-component solutions were reviewed again to identify the point at which all poorly performing items had been removed, and all item loadings met the target of 'good' (≥0.55). However, given that item loadings were generally 'good', a higher target of 'very good' (>0.63) was set in order to retain the smallest set of items. The variance explained by each solution was also considered. In the first iteration of the 3-component solution (Table 9.12), all but one item (10.22) had 'good' loadings and the solution explained 71.14% of the variance. In rotated PCA 17 (Table 9.21), the solution explained 75.10% of the variance and all item loadings were 'very good'. However, in the previous solution (Table 9.20), which explained 75.14% of the variance, 'Have you had to work around your RA stiffness (or do things in a different way)?' had been identified for removal as it had the lowest loading. The removal of 'Have you had to work around your RA stiffness (or do things in a different way)?' had led to a reduction in size of the smallest component ("Psychosocial"). As the loading of 'Have you had to work around your RA stiffness (or do things in a different way)?' (.620) was very close to the 'very good' cut-off and retaining it would preserve the number of items in an already small component, rotated PCA 16 (Table 9.20), with 23 items, was identified as the most appropriate 3-component solution from the combined perspective of component loadings and explained variance.

Having identified which 3-component solution to retain, an attempt was made to enhance it by testing the removal of similar items and trying to even up the size of the components. Firstly, the two items capturing MS severity ('How would you describe the overall level of morning stiffness you have had from the time you wake up?' (NRS and 5-option ordinal scale)) were investigated. When reviewing the retained 3-component solution (Table 9.20) the component loadings of 'How would you describe

the overall level of morning stiffness you have had from the time you wake up?' on both response scales (NRS and 5-option ordinal scale) were 'excellent', although the loading of the NRS item (0.750) was very slightly higher than the loading of the 5-option ordinal scale item (0.749). When looking at previous 3-component solutions, the NRS item loaded higher than the 5-option ordinal scale item on more occasions (n=6, n=3). Furthermore, when looking at the loadings of items within the individual components (Section 9.4.2.4.2), the NRS item loaded higher (0.925) than the 5-option ordinal scale item (0.895). As a result, the 5-option ordinal scale item was removed.

The resulting solution (Table 9.29) included 22-items, retained the 3-component structure, and explained 75.04% of the variance. All item loadings were 'very good' and the Cronbach's alpha coefficient was 0.955. Next, the two items capturing stiffness severity ('Please circle the number that best describes the severity of your RA stiffness over the past 7 days' and 'Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week') were explored. In rotated PCA 16 (Table 9.20) the loadings of both items were 'excellent', although the loading of 'Please circle the number that best describes the severity of your RA stiffness over the past 7 days' (0.765) was higher than for 'Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week' (0.761). When looking back at the loadings of these items in the 3-component solutions, 'Please circle the number that best describes the severity of your RA stiffness over the past 7 days' loaded higher than 'Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week' on all occasions (n=9). Furthermore, when looking at the loadings of items within the individual components (Section 9.4.2.4.2), it was again found that 'Please circle the number that best describes the severity of your RA stiffness over the past 7 days' loaded slightly higher (0.951) than 'Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week' (0.946). As a result, 'Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week' was removed.

Table 9.29: Rotated PCA ('How would you describe the overall level of morning stiffness you have had from the time you wake up?' (5-option ordinal scale) removed)

Item	1	2	3
Draft item severity	**.750**		.446
Traditional item severity G	**.744**		.436
Draft item timing	**.743**		
Traditional item severity A	**.725**	.424	
Draft item in joints	**.719**		
Draft item impact	**.697**		.518
Draft item importance	**.661**		.569
Draft item after immobility	**.650**		
Traditional item duration D	**.641**		
Draft item wash		**.823**	
Draft item dress		**.775**	
Draft item responsibilities		**.714**	
Draft item daily tasks	.460	**.696**	
Draft item get out of bed		**.671**	
Draft item grip		**.658**	
Draft item fine movement		**.648**	
Draft item balance		**.644**	
Draft item worried			**.809**
Draft item self-conscious			**.770**
Draft item frustrated			**.708**
Draft item change plans		.487	**.700**
Draft item work around		.572	**.622**
Component key: Physical; Severity; Psychosocial			

The resulting solution (Table 9.30) included 21-items and retained the 3-component structure (although the "Severity" and "Physical" switched places). The solution explained 74.53% of the variance. All item loadings were 'very good' and the Cronbach's alpha coefficient was .951. The components were also slightly more equal in size ("Severity" and "Physical" n=8, "Psychosocial" n=5). The solution did demonstrate seven items which loaded >.40 on more than one component however this is fewer than in previous solutions such as rotated PCA 16 (Table 9.20) which had 12 multiple loading items. This solution provided a good performing, well balanced, conceptually sound set of items that was felt to be an improvement on that identified by statistics alone. Therefore it was taken forward as the best structure for the provisional RA stiffness PROM for further testing. As before, in subsequent reference to stiffness items, the full item wording is used in the text while abbreviated item wording is used in tables but in both, 'draft' and 'traditional' have been replaced with 'final'.

Table 9.30: Rotated PCA ('Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week') removed)

Item	1	2	3
Draft item wash	**.826**		
Draft item dress	**.776**		
Draft item responsibilities	**.719**		
Draft item daily tasks	**.703**	.447	
Draft item get out of bed	**.668**		
Draft item balance	**.664**		
Draft item grip	**.660**		
Draft item fine movement	**.645**		
Draft item timing		**.763**	
Draft item in joints		**.734**	
Draft item severity		**.716**	.458
Traditional item severity A	.441	**.698**	
Draft item impact		**.668**	.531
Draft item after immobility		**.664**	
Traditional item duration D		**.632**	
Draft item importance		**.631**	.581
Draft item worried			**.815**
Draft item self-conscious			**.775**
Draft item frustrated			**.716**
Draft item change plans	.493		**.702**
Draft item work around	.574		**.624**

Component key: Physical; Severity; Psychosocial

9.4.2.5 Bootstrapping

The most appropriate solution (Table 9.30) was then subject to bootstrapping to test whether the components within the structure were retained during repeated PCA of subsets of the original data, indicating stable components. Twenty rounds of rotated PCA were performed, each time with randomly selected subsets of 50% of the dataset (Table 9.31). On five occasions bootstrapping produced a 2-component solution similar to that identified in rotated PCA 21 (Table 9.25). However, on 15 occasions, a 3-component solution was produced indicating good stability. On seven occasions the solution was identical to the final solution. During the remaining eight rounds of bootstrapping, there was some movement of items between components, as a result of items which loaded >0.40 on more than one component. This meant that the number of items in some components did change however, the component structure was retained. This demonstrated the strength of the individual components, particularly "Severity" and "Physical" which were slightly larger than "Psychosocial".

Table 9.31: Component structure and item placement during rotated PCA on 20 randomly selected subsets of the dataset

Round	FS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
No. of components	3	3	2	3	3	3	2	3	3	3	3	3	2	3	3	3	3	2	2	3	3
Final item timing	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Ph	Se	Se
Final item in joints	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Ph	Se	Se
Final item after immobility	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Ph	Se	Se
Final item dress	Ph	Ph	Ph	Ph	Ph	Ph	Se	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph
Final item wash	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph
Final item responsibilities	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph
Final item daily tasks	Ph	Ph	Ph	Ph	Ph	Ph	Se	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph
Final item get out of bed	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph
Final item fine movement	Ph	Ph	Se	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Se	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph
Final item grip	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Se	Ph	Ph	Ph
Final item balance	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph	Ph
Final item frustrated	Ps	Ps	Ph	Ps	Ps	Ps	Ph	Ps	Ps	Ps	Ps	Ps	Ph	Ps	Ps	Ps	Ps	Ph	Ps	Ps	Ps
Final item worried	Ps	Ps	Ph	Ps	Ps	Ps	Ph	Ps	Ps	Ps	Ps	Ps	Ph	Ps	Ps	Ps	Ps	Ph	Ps	Ps	Ps
Final item self-conscious	Ps	Ps	Ph	Ps	Ps	Ps	Ph	Ps	Ps	Ps	Ps	Ps	Ph	Ps	Ps	Ps	Ps	Ph	Ps	Ps	Ps
Final item change plans	Ps	Ph	Ph	Ps	Ps	Ps	Ph	Ps	Ps	Ps	Ps	Ps	Ph	Ps	Ps	Ph	Ps	Ph	Ps	Ps	Ps
Final item work around	Ps	Ph	Ph	Ph	Ps	Ps	Ph	Ph	Ps	Ps	Ps	Ps	Ph	Ps	Ph	Ph	Ps	Ph	Ps	Ps	Ph
Final item impact	Se	Se	Se	Se	Se	Se	Se	Se	Se	Ps	Se	Se	Se	Se	Se	Se	Se	Se	Ps	Se	Se
Final item severity	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Ph	Se	Se
Final item importance	Se	Se	Se	Ps	Se	Se	Ph	Ps	Se	Ps	Se	Se	Se	Se	Se	Se	Ps	Ph	Ps	Se	Ps
Final item severity A	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Ph	Se	Se
Final item duration D	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Se	Ps	Se	Ps

FS = Final solution; Ph = Physical component; Se = Severity component; Ps = Psychosocial component

9.4.2.6 Scoring

All items within "Physical" and "Psychosocial" were assessed on the same scale while items within "Severity" were assessed on different scales. Scores using simple summation are shown in Table 9.32. Factor scores were also generated using the regression method in SPSS. Scores were generated for all participants for the "Severity" and "Psychosocial" and 276 scores for "Physical" because one participant provided no responses to any item in that component.

Table 9.32: Each component and the scale each item was scored on

Physical component		**Severity component**		**Psychosocial component**	
Item	**Score**	**Item**	**Score**	**Item**	**Score**
Final item wash	0-3	Final item timing	0-4	Final item worried	0-3
Final item dress	0-3	Final item in joints	0-3	Final item self-conscious	0-3
Final item responsibilities	0-3	Final item severity	0-0	Final item frustrated	0-3
Final item daily tasks	0-3	Final item severity A	0-10	Final item change plans	0-3
Final item get out of bed	0-3	Final item impact	0-10	Final item work around	0-3
Final item balance	0-3	Final item after immobility	0-3		
Final item grip	0-3	Final item duration D	0-6		
Final item fine movement	0-3	Final item importance	0-10		
Range	0-24	**Range**	0-56	**Range**	0-15
Total possible range of score=0-95					

To compare the simple summation and factor score methods, correlations (Table 9.33) and scatterplots (Figures 9.1-9.3) were produced comparing each participant's component simple summation scores with their factor scores.

Table 9.33: Correlation between simple summation and regression factor scores for each component

	Physical component (F1) sum	**Severity component (F2) sum**	**Psychosocial component (F3) sum**
REGR factor score F1	.784	.394	.460
REGR factor score F2	.445	.760	.377
REGR factor score F3	.436	.523	.809
All significant at $p<.01$			

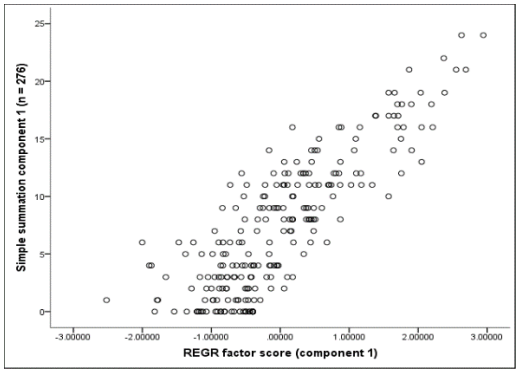


Figure 9.1: "Physical" sum score plotted against "Physical" regression factor score

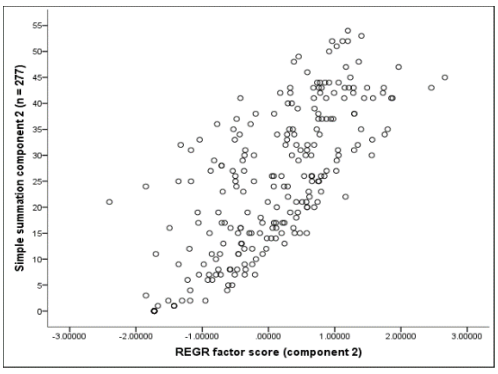


Figure 9.2: "Severity" sum plotted against "Severity" regression factor score

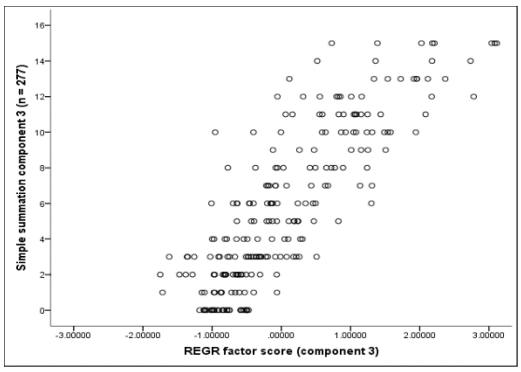


Figure 9.3: "Psychosocial" sum plotted against "Psychosocial" regression factor score

As would be expected, the correlations between the scores generated using the simple summation method and the regression factor scores were strong between the same components, and weak or moderate between different components. The scores generated using the simple summation method were significantly related to those generated using the regression factor score method (Table 9.33). The strong correlation between the scores suggested similarities in approaches. The surprising effectiveness of simple summation has been demonstrated in the literature (Dawes, 1979), and has the advantage of ease of application in clinical or research environments. The factor score method produces a more accurate indication of the relationship between each item and the component (Field, 2009). However, it would be more complicated to implement in applied environments.

9.4.2.6.1 Whole scale or individual component scoring

Discussion with the supervisory team regarding the scale scoring highlighted two areas for consideration. Firstly, given the identification of three individual components, the principal of a total score was questioned in favor of scores for each individual component. However, regardless of the most appropriate scoring of the scale it was still considered important to provide the option to generate a sum of all components. Secondly, it was highlighted that neither scoring approach represented the PCA output. As the range of possible scores for "Severity" was much larger (0-56) than the other components ("Physical"=0-24; "Psychosocial"=0-15), it may appear that it was more important and its influence dominated the total score calculation. Therefore, to ensure that all components were comparable, and that generation of a sum of all components was possible, item scores were rescaled using percentages (Tables 9.34-9.36). A percentage score for each individual component could then be generated, which could be used to create a total percentage score for the new RA stiffness PROM by adding together the percentage scores for each individual component and dividing them by three. This approach provided balanced scores across components. The scoring protocol can be viewed in Appendix Z.

Table 9.34: Rescaled percentage score for "Physical"

Item	Original score	Percentage score
Final item wash		
Final item dress		
Final item responsibilities		0=0%
Final item daily tasks	0-3	1=33%
Final item get out of bed		2=67%
Final item balance		3=100%
Final item grip		
Final item fine movement		

Table 9.35: Rescaled percentage score for "Severity"

Item	Original score	Percentage score
Final item in joints Final item after immobility	0-3	0=0% 1=33% 2=67% 3=100%
Final item timing	0-4	0=0% 1=25% 2=50% 3=75% 4=100%
Final item duration D	0-6	0=0% 1=17% 2=33% 3=50% 4=67% 5=83% 6=100%
Final item severity A Final item impact Final item severity Final item importance	0-10	0=0% 1=10% 2=20% 3=30% 4=40% 5=50% 6=60% 7=70% 8=80% 9=90% 10=100%

Table 9.36: Rescaled percentage score for "Psychosocial"

Item	Original score	Percentage score
Final item worried Final item self-conscious Final item frustrated Final item change plans Final item work around	0-3	0=0% 1=33% 2=67% 3=100%

9.4.2.6.2 Dealing with missing data

Simple summation and factor scores had been generated with all available data regardless of the amount of missing data. However, for use in applied situations a more accurate approach was required. Similar to guidance provided in other scales such as the HAQ (Fries *et al,* 1980), it was decided that when scoring the scale, one missing item per component was acceptable. Here, individual component percentage scores could be generated by adding together the item percentages from the available data and dividing that by the number of item percentages provided. A score for each

individual component would be required for the generation of a sum of all components. Future work investigating the treatment of missing data using other data sets is required.

In this dataset, six participants had more than one missing item on some components therefore, scores were not calculated for these participants for these components ("Physical" n=272, "Severity" n=274, "Psychosocial" n=277). A sum of all components was only calculated for participants with a score for each component (n=271).

9.4.2.6.3 Frequency and distribution of the components and total score

Frequency and distributions of each individual component and the sum of all components were generated based on the percentage scores described above (Figures 9.4-9.7).

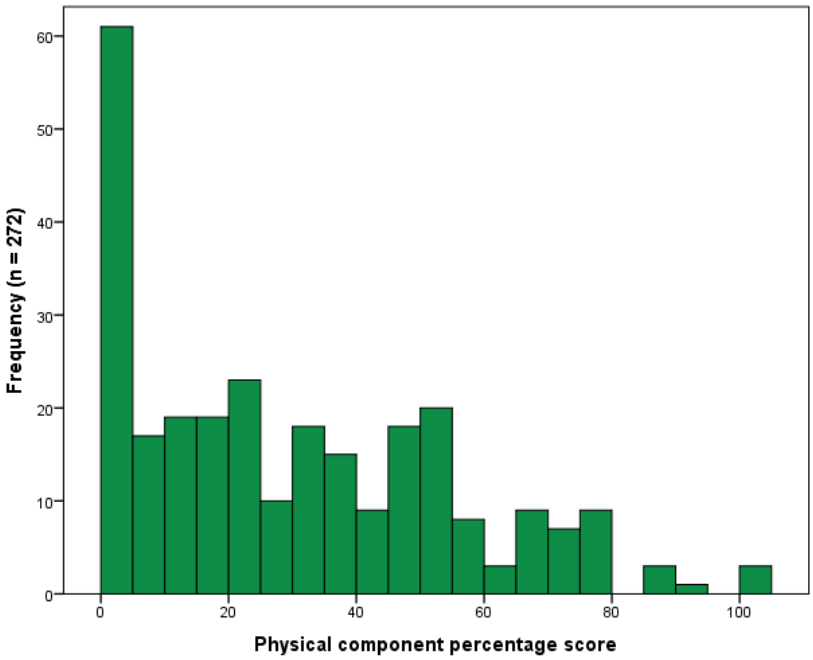


Figure 9.4: Frequency and distribution for "Physical"

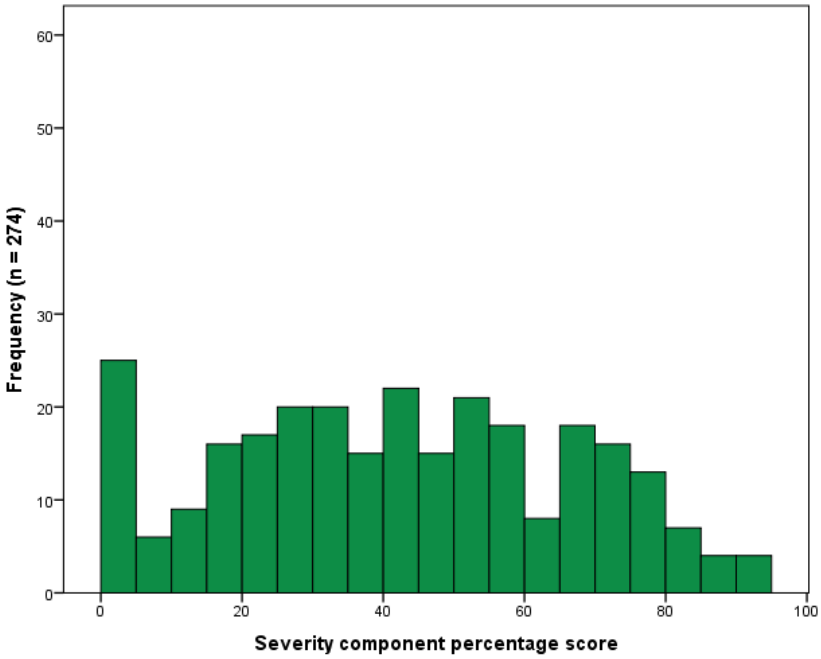


Figure 9.5: Frequency and distribution for "Severity"

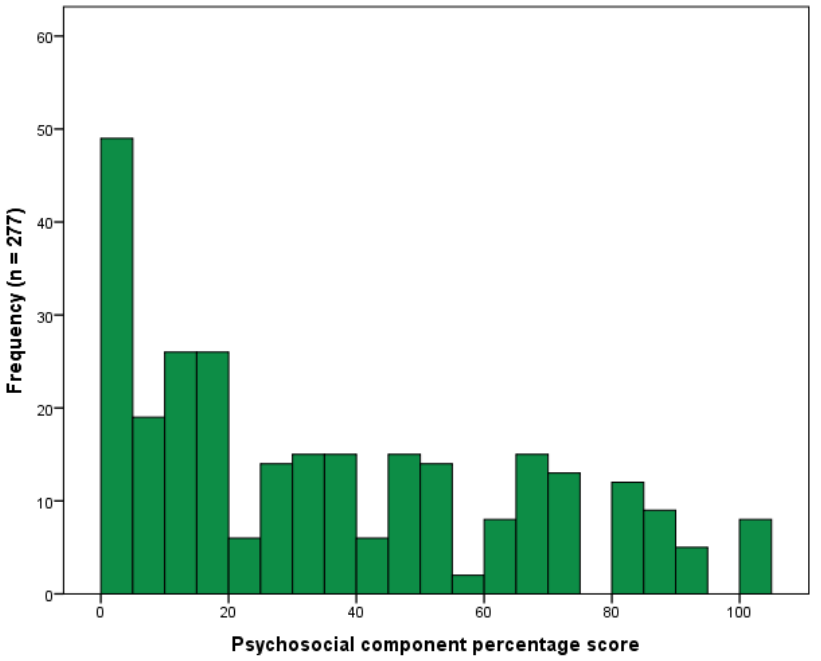


Figure 9.6: Frequency and distribution for "Psychosocial"

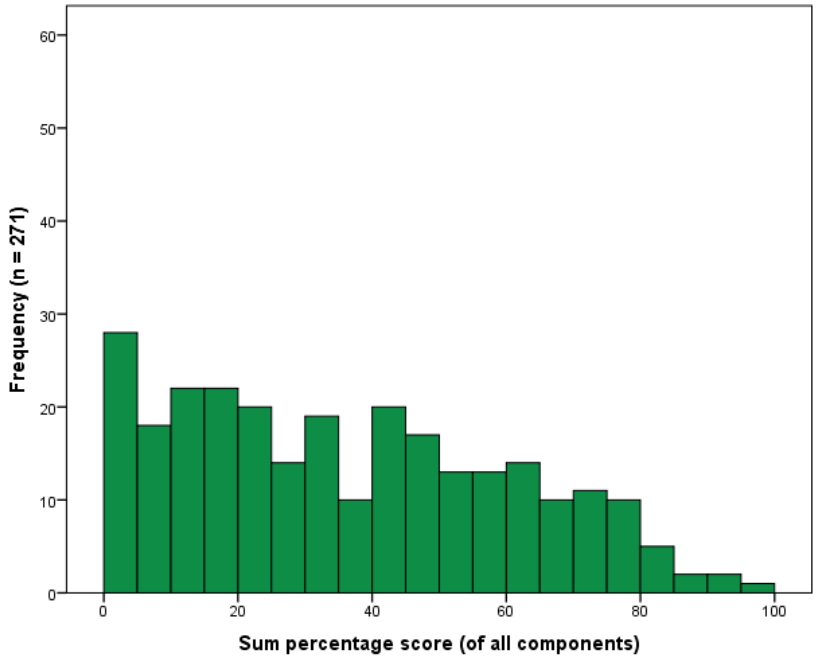


Figure 9.7: Frequency and distribution for the sum of all components

Frequency and distributions indicated that participants reported a range of scores on each component and the sum of all components, although there were higher frequencies at the lower end of each.

9.4.2.7 Formatting

Component names used thus far ("Physical", "Severity", and "Psychosocial") had been developed based on consideration of their content following discussion with the supervisory team. Discussion with one patient partner (GB) reinforced the appropriateness of these component names from the patient perspective. As recommended by GB, a summary sentence that captured the content of each component was added for clarity. Items within each component were presented in individually marked boxes. This layout ensured that the individual components could retain their distinct nature. Discussion of the final PROM items with one member of the supervisory team (JK) highlighted the importance of item placement specifically in relation to two items ('How would you describe the overall level of morning stiffness you have had from the time you wake up?' and 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?') which had different timeframes to other items within the questionnaire. Rational placement of these items was important in order to retain them as tested during the survey. The order that the items were tested in was retained where possible. The layout and font of the final PROM was also discussed with one patient partner (GB) to ensure appropriateness from the

patient perspective. The final layout, now called the RA stiffness (RAST) PROM is detailed in Appendix AA.

9.4.3 Stage 4 analysis: Validity testing

The clinical data captured in the survey, but so far not used in the development of the RAST, provided the opportunity for preliminary validity testing. Spearman's rank order correlations were used to investigate whether the RAST demonstrated expected relationships with other measures of RA. Correlation coefficients between the RAST (each individual component and total percentage scores) and clinical measures capturing disease activity (PDAS2, Choy *et al,* 2008; Choy *et al,* 2015), disability (MHAQ, Pincus *et al,* 1983), pain (Pain NRS, Farrah *et al,* 2001; Hawker *et al,* 2011), fatigue (BRAF severity-NRS, Nicklin *et al,* 2010a; Nicklin *et al,* 2010b), patient global assessment (PtG VAS, van der Heijde *et al,* 1993) and current flare (PFQ, Bykerk *et al,* 2012; Bykerk *et al,* 2014b) were calculated and shared variance reported. The PDAS2 (Choy *et al,* 2008; Choy *et al,* 2015) without EMS (Choy and Leung, 2016) was used here to avoid circular reasoning.

Given that there has been little robust investigation to date into stiffness assessment, Spearman's rank order correlations and shared variance were also used to investigate the relationships between traditional stiffness severity ('How would you describe the overall level of morning stiffness you have had from the time you wake up?' on NRS and 5-option ordinal scale and 'Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week') and duration ('Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?' and 'How long does your morning stiffness last from waking until maximum improvement occurs?' on 3-option ordinal scale and in minutes and hours) items and other measures of RA (as above).

9.4.4 Stage 4 results: Validity testing

Descriptive statistics for the non-stiffness items were reported during description of the study population (Chapter 8) and histograms and investigation into normality for these data are available in Appendix W.

9.4.4.1 Comparison between RAST and clinical measures

Table 9.37 reports the correlation coefficients (*r*) and shared variance (R^2) and demonstrated strong correlations with disease activity and moderate to strong

correlations with other measures of disease, apart from current flare which demonstrated moderate or weak correlations.

Correlations between RAST and other measures of disease were higher than the expected moderate correlations. However, these appeared appropriate when considering the pattern across components, which also provided support for the 3-component structure. “Physical” shared the most variance with disability (r_S=0.886, R^2=78.5%), which may be expected given the overlap in concepts and the response option format with the MHAQ (Pincus *et al,* 1983). “Severity” shared the most variance with pain (r_S=0.851, R^2=72.4%) and PtG (r_S=0.826, R^2=68.2%), which would be expected and may in part be due to similarities in assessment format (e.g. NRS or VAS). “Psychosocial” shared less variance with the disease related measures than other components. This may indicate that it is capturing aspects not currently being assessed (Pincus *et al,* 1983).

When considering these results specifically in terms of shared variance, while RAST (individual components and sum score) appears related to other measures of disease, it does not appear to be capturing the same information. Shared variance between RAST (individual components and sum score) and pain (R^2=44.9-72.4%) and fatigue (R^2=43.2-55.2%) varied depending on the component. The relationship between RAST (individual components and sum score) and the PtG (R^2=48.2-68.2%) may be expected given the suggestion that the PtG provides a patient report of the impact of RA rather than disease activity (Kalyoncu *et al,* 2009; French *et al,* 2013). Therefore, the shared concept of impact may increase this relationship. The relatively weak relationship between RAST (individual components and sum score) and self-reported current flare (r_S=0.455-0.532, R^2=20.7-28.3%) was unexpected and is considered further in the discussion.

Table 9.37: Correlation coefficients (*r*) and shared variance (R^2) between RAST (components and sum) and measures of disease

		Physical (1)	Severity (2)	Psycho-social (3)	Sum	Disease activity (DA)	Disability (Dis)	Pain	Fatigue	PtG	Flare
1	*r* (n)	-	.835 (271)	.836 (272)	.942 (271)	.829 (272)	.886 (272)	.752 (272)	.699 (272)	.764 (270)	.455 (271)
	R^2 (%)	-	69.7%	69.7%	88.7%	68.7%	78.5%	56.6%	48.9%	58.4%	20.7%
2	*r* (n)		-	.804 (274)	.927 (271)	.882 (274)	.766 (274)	.851 (274)	.741 (274)	.826 (272)	.532 (273)
	R^2 (%)		-	64.6%	85.9%	77.8%	58.7%	72.4%	54.9%	68.2%	28.3%
3	*r* (n)			-	.945 (271)	.746 (277)	.744 (277)	.670 (277)	.657 (277)	.694 (275)	.458 (276)
	R^2 (%)			-	89.3%	55.7%	55.4%	44.9%	43.2%	48.2%	21.0%
Sum	*r* (n)				-	.868 (271)	.848 (271)	.810 (271)	.743 (271)	.804 (269)	.521 (270)
	R^2 (%)				-	75.3%	71.9%	65.6%	55.2%	64.6%	27.1%
DA	*r* (n)					-	.792 (277)	.617 (276)	.609 (276)	.596 (274)	.444 (275)
	R^2 (%)					-	62.7%	38.1%	37.1%	35.5%	19.7%
Dis	*r* (n)						-	713 (277)	.646 (277)	.728 (275)	.369 (276)
	R^2 (%)						-	50.8%	41.7%	53.0%	13.6%
Pain	*r* (n)							-	.693 (277)	.827 (275)	.576 (276)
	R^2 (%)							-	48.0%	68.4%	33.2%
Fatigue	*r* (n)								-	.700 (275)	.445 (276)
	R^2 (%)								-	49.0%	19.8%
PtG	*r* (n)									-	.511 (274)
	R^2 (%)									-	26.1%
Flare	*r* (n)										-
	R^2 (%)										-

All significant at $p<.01$; Physical (1)=Physical component % score; Severity (2)=Severity component % score; Psychosocial (3)=Psychosocial component % score; DA=Disease activity (PDAS2, Choy *et al,* 2008; Choy *et al,* 2015; Choy and Leung, 2016); Dis.=Disability (MHAQ, Pincus *et al,* 1983); Fat.=Fatigue (BRAF severity-NRS, Nicklin *et al,* 2010a; Nicklin *et al,* 2010b)

9.4.4.2 Comparison between RAST and disease activity measures

Strong correlations and a range of shared variance (r_s=0.746-.882, R^2=55.7-77.8%) were reported between RAST (individual components and sum score) and disease activity suggesting that RAST is assessing an aspect of patient reported disease activity as assessed by the PDAS2 (Choy *et al,* 2008; Choy *et al,* 2015). Interestingly, SJC individually only explained 31.4-41.5% of the variance in RAST (individual components and sum score). This indicates that that RAST may capture something not currently included within this aspect of disease activity assessment. It also reflects earlier qualitative work where some patients reported that inflammation was independent of stiffness and some stated they were unable to recognise inflammation (Section 4.5.2.1.3).

9.4.4.3 Comparison between traditional stiffness items and clinical measures

Traditional stiffness items demonstrated moderate to strong correlations with disease activity and weak to strong correlations with other measures of disease (Table 9.38). As demonstrated in the literature, generally stronger correlations were reported for traditional severity items than for traditional duration items (Section 2.4.1). Traditional item duration D ('Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?') demonstrated moderate correlations with all measures of disease apart from flare, whilst traditional items duration E and F ('How long does your morning stiffness last from waking until maximum improvement occurs?' on 3-option ordinal scale and in minutes and hours) only demonstrated weak to moderate correlations. Similarly when considering the relationship between traditional stiffness items and disease activity, all traditional severity items demonstrated strong correlations with disease activity, while the traditional duration items ('How long does your morning stiffness last from waking until maximum improvement occurs?' on 3-option ordinal scale and in minutes and hours) or strong ('Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?') demonstrated moderate correlations with disease activity. When considered another way, traditional items duration E and F ('How long does your morning stiffness last from waking until maximum improvement occurs?' on 3-option ordinal scale and in minutes and hours) only shared a quarter of the variance in disease activity while all other items shared over half.

Furthermore, when looking at correlations between items, whilst traditional severity items correlated strongly with each other and traditional duration items correlated strongly with each other, traditional severity and duration items only correlated

moderately with each other. This reinforces that these concepts may capture different information and supports the inclusion of both within RAST. In contrast, RAST demonstrated strong internal correlations between individual components and the sum score (Table 9.37).

Table 9.38: Correlation coefficients (*r*) and shared variance (R^2) between traditional stiffness items and clinical measures

		A	G	C	D	E	F	DA	Dis.	Pain	Fat.	PtG	Flare
A	*r* (n)	-	.888 (273)	.864 (274)	.718 (273)	.497 (237)	.564 (219)	.830 (274)	.729 (274)	.814 (274)	.718 (274)	.798 (272)	.441 (273)
	R^2 (%)	-	78.9%	74.6%	51.6%	24.7%	31.8%	68.9%	53.1%	66.3%	51.6%	63.7%	19.4%
G	*r* (n)		-	.844 (275)	.706 (274)	.536 (237)	.552 (221)	.850 (275)	.721 (275)	.858 (275)	.707 (275)	.806 (273)	.508 (274)
	R^2 (%)		-	71.2%	49.8%	28.7%	30.5%	72.3%	52.0%	73.6%	50.0%	65.0%	25.8%
C	*r* (n)			-	.706 (275)	.527 (238)	.564 (221)	.784 (276)	.667 (276)	.770 (276)	.654 (276)	.745 (274)	.475 (275)
	R^2 (%)			-	49.8%	27.8%	31.8%	61.5%	44.5%	59.3%	42.8%	55.5%	22.6%
D	*r* (n)				-	.699 (237)	.753 (237)	.726 (276)	.595 (276)	.656 (276)	.628 (276)	.641 (274)	.462 (275)
	R^2 (%)				-	44.8%	56.7%	52.7%	35.4%	43.0%	39.4%	41.1%	21.3%
E	*r* (n)					-	.779 (211)	.508 (238)	.371 (238)	.543 (238)	.435 (238)	.516 (237)	.354 (237)
	R^2 (%)					-	63.8%	25.8%	13.8%	29.5%	18.9%	26.6%	12.5%
F	*r* (n)						-	.551 (221)	.424 (221)	.521 (221)	.458 (221)	.552 (221)	.333 (220)
	R^2 (%)						-	30.4%	18.0%	27.1%	21.0%	30.6%	11.1%
DA	*r* (n)							-	.792 (277)	.812 (277)	.706 (277)	.824 (275)	.532 (276)
	R^2 (%)							-	62.7%	65.9%	49.8%	67.9%	28.3%
Dis.	*r* (n)								-	.713 (277)	.646 (277)	.728 (257)	.369 (276)
	R^2 (%)								-	50.8%	41.7%	53.0%	13.6%
Pain	*r* (n)									-	.693 (277)	.827 (275)	.576 (276)
	R^2 (%)									-		68.4%	33.2%
Fat.	*r* (n)										-	.700 (275)	.445 (276)
	R^2 (%)										-		19.8%
PtG	*r* (n)											-	.511 (274)
	R^2 (%)											-	26.1%
Flare	*r* (n)												-
	R^2 (%)												-

All significant at $p<.01$; A=Traditional item severity A (MS severity NRS); G=Traditional item severity G (stiffness severity NRS); C=Traditional item severity C (MS severity 5-point Likert scale); D=Traditional item duration D (MS duration 6 ordinal response options); E=Traditional item duration E (MS duration 3 ordinal response options); F=Traditional item duration F (MS duration minutes/hours); DA=Disease activity (PDAS2, Choy *et al,* 2008; Choy *et al,* 2015; Choy and Leung, 2016); Dis.=Disability (MHAQ, Pincus *et al,* 1983); Fat.=Fatigue (BRAF severity-NRS, Nicklin *et al,* 2010a; Nicklin *et al,* 2010b)

9.5 Discussion

A rigorous examination of the draft stiffness items in relation to component structure and internal consistency, and from statistical and conceptual perspectives, led to the specification of the smallest and most effective combination of items reflecting the patients' experience of stiffness. This enabled the development of a proposed new RA stiffness PROM (RAST) containing 21-items across three components ("Physical", "Severity", and "Psychosocial") which reflect the patient experience of stiffness. During preliminary validity testing, the RAST demonstrated stronger correlations for every variable, apart from PtG where very similar strong correlations were demonstrated (Table 9.37 and 9.38). Furthermore, its rigorous item development process, consistent with PROM development guidelines (USDHHS FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b), provides superior face and content validity in comparison to traditional stiffness items and is a novel characteristic of the new RAST.

In relation to the conceptual development of the RAST, the final content and structure reflects the patient experience of stiffness identified in earlier qualitative studies (Chapter 4 and 5). Most items in RAST were based on stiffness over a seven day timeframe. However, the inclusion of traditional stiffness items ('How would you describe the overall level of morning stiffness you have had from the time you wake up?' and 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?') in the final RAST resulted in items with different timeframes. 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?' asks about stiffness today and 'How would you describe the overall level of morning stiffness you have had from the time you wake up?' did not specify a timeframe. It is proposed that as both traditional stiffness items ('How would you describe the overall level of morning stiffness you have had from the time you wake up?' and 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?') related to MS specifically, short timeframes such as 'today' may be more appropriate, especially given the variability in the experience (Chapters 4 and 5). Furthermore, it was proposed that as both 'Please circle the number that best describes the severity of your RA stiffness over the past 7 days' and 'How would you describe the overall level of morning stiffness you have had from the time you wake up?' captured severity but over different timeframes, it may be that the combination of both items captures stiffness variability (e.g. a low score on 'Please circle the number that best describes the severity of your RA stiffness over the past 7 days' and a high score on 'How would

you describe the overall level of morning stiffness you have had from the time you wake up?' could suggest that stiffness severity has been low over the last 7 days but is severe today). The concept of variability was originally identified as relevant in the patient experience of stiffness (Chapter 4) and was captured in an item (Chapter 6), but was later removed for poor performance (Chapter 8). However, further qualitative work is required to fully understand the concept and timeframe that 'How would you describe the overall level of morning stiffness you have had from the time you wake up?' is capturing. The variation in timeframes across items also had implications on the placement of the items in the final questionnaire to ensure that the neutral timeframe was retained.

RAST includes three aspects of the impact triad (Sanderson *et al,* 2011), severity, importance and impact, but not coping. 'Please circle the number that best describes how well you have coped with your RA stiffness over the past 7 days' demonstrated poor item loadings in the initial exploration of the component structure (Chapter 8) and performed poorly during investigations into the whole and individual component structure, and was removed in the first round of item reduction testing. As discussed during item development (Chapter 6) the impact triad items were developed based on previous work (Nicklin, 2009; Nicklin *et al,* 2010a; Nicklin *et al,* 2010b). Specifically in relation to the coping item, a reversed anchor layout had been suggested as being more appropriate by patients (Nicklin, 2009), and was implemented here. Although the inconsistency of the anchors in comparison to other items was highlighted by some participants in cognitive interviews (Chapter 7), the feedback was not all negative and it was felt that the anchor format should be retained given the evidence in support of it from previous work (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b). However, on review of responses to this item during the survey it appeared likely that some participants had marked the opposite response to that intended, although it was not possible to tell which patients marked responses inadvertently and those who did it intentionally. It is likely that this affected the correlations between this item and other items and was a contributing factor to its exclusion. A very recent study investigated the use of cognitive interviews in the translation of the BRAF (Nicklin *et al,* 2010a; Nicklin *et al,* 2010b) and RAID (Gossec *et al,* 2009; Gossec *et al,* 2011) questionnaires into six European languages (Hewlett *et al,* 2016a). The study found that there were difficulties with capturing and interpreting coping. It also highlighted problems with the anchor placement in the BRAF coping question as seven of the 10 Dutch participants marked the item in the opposite way to that verbally conveyed. To investigate this further, a survey study was performed including both the original and

a revised coping item, which found that the revised item with anchors in the traditional place performed better (Hewlett *et al*, 2016b). This provides evidence to support the suggestion that responses to the coping item in this study were influenced by the anchor placement. Further work into the concept of coping and how to assess it may have implications for the RAST.

A final consideration relating to RAST is the appropriateness of a sum score. To reflect the three components identified during PCA, a score was generated for each individual component that could be used for different purposes. If the purpose was to assess the severity of stiffness then "Severity" could be used. If the purpose was to assess specific types of stiffness impact then the respective impact components ("Physical" and "Psychosocial") could be employed. It was recognised that in some circumstances, all components would be used, therefore the option was provided to enable a sum score of all components if required.

As reported earlier (Section 2.4), there is currently no clear evidence regarding the most appropriate measure to use to assess stiffness in RA. This work provides the first comprehensive and robust evaluation of stiffness assessment involving the RAST and traditional stiffness items. One key finding was the relationship between stiffness and disease activity where RAST (individual components and sum score) demonstrated strong correlations. Strong correlations were also reported for traditional severity items while moderate or strong correlations were reported for traditional duration items. Correlations between stiffness items and disease activity assessed using the PDAS2 (Choy *et al,* 2008) have not been demonstrated before. Although MS is traditionally considered an indicator of inflammatory activity in RA (e.g. Lansbury, 1956; Hazes *et al,* 1994; Soubrier *et al,* 2006), there is little evidence of this relationship in the literature. Two previous studies specifically exploring the relationship between stiffness and disease activity demonstrated weak or moderate correlations (Westhoff *et al,* 2008; Khan *et al,* 2009). Khan *et al.* (2009) reported weak correlations between MS duration and DAS28 (r_S=0.46, p<0.001) while Westhoff *et al.* (2008) reported weak or moderate correlations between DAS28 and MS severity (baseline r_S=0.47, follow up r_S=0.58, both p<0.001). The MS duration results are similar to results from this study where 'How long does your morning stiffness last from waking until maximum improvement occurs?' reported in minutes and hours (the same item as used by Khan *et al.* (2009)) correlated only moderately with disease activity (r_S=0.508, p<0.01). However, despite similarities, disease activity was assessed differently. The results for MS severity identified by Westhoff *et al.* (2008)

were much lower than the strong correlations identified in this study. However, stiffness items were not directly comparable across studies.

The consideration of different disease activity assessment is important. This study assessed disease activity using the PDAS2 without EMS which is a validated patient report of disease activity (Choy *et al,* 2008; Choy *et al,* 2015). Although the PDAS2 demonstrated strong correlations with the DAS28 (r_S=0.76) during validation (Choy *et al,* 2008), the DAS28 is still the recommended tool for assessment of disease activity (e.g. Luqmani *et al,* 2009). Research suggests that self-reported measures are more closely associated with other self-reported measures than with laboratory or physician-reported measures, specifically in relation to stiffness (Khan *et al,* 2009; Westhoff *et al,* 2008), but also more broadly in RA (Pincus *et al,* 1989; Taal *et al,* 1998). In this study, the self-reported nature of all data may have influenced the demonstrated relationship between disease activity and stiffness. An important area for further testing of RAST includes comparison with other disease activity assessments that contain objective items, such as the DAS28 and blood tests for inflammatory markers.

Another important result was the demonstration of the poor performance of traditional stiffness duration items ('How long does your morning stiffness last from waking until maximum improvement occurs?' on 3-option ordinal scale and in minutes and hours). Stiffness assessment in research trials is most common in the form of MS duration (Kalyoncu *et al,* 2009), yet there are difficulties with its assessment (e.g. Vliet Vlieland *et al,* 1997), and some literature suggests that stiffness severity items have better measurement properties than duration items (e.g. Lie *et al,* 2014) (Section 2.4.1.4). One previous study reported that despite assessing both MS severity and duration, only MS severity was used in analyses because it was more responsive and MS duration had a high proportion of missing data (Westhoff et al, 2008). This was consistent with the findings in the present study where 'How long does your morning stiffness last from waking until maximum improvement occurs?' on 3-option ordinal scale and in minutes and hours had large amounts of missing data (Section 8.4.2.3) and were excluded from PROM development analyses. Some was a result of inaccurate item completion where survey participants marked the minutes or hours response options provided rather than stating a specific duration. One participant during the cognitive interviews (Chapter 7) had suggested that item 12.5 was not clear enough that an amount of time had to be specified (*"minutes, hours, that could be hours, okay, not minutes" [2408]*). When reviewing 'How long does your morning

stiffness last from waking until maximum improvement occurs?' on 3-option ordinal scale and in minutes and hours with the supervisory team, it was discussed that these items do not provide clear options to respond 'none' or 'no stiffness'. This was proposed as a reason for the large amounts of missing data and thus poor correlations demonstrated by these items. It may also explain why tradition duration item 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?' (which provides an option for 'no stiffness') demonstrated higher amounts of complete data, and superior correlations than the two other duration items. As the concept of duration had been challenged by patients (Chapter 5), it was interesting that traditional stiffness duration item 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?' was retained as an item in the final RAST. It may be that because duration is so commonly used to assess stiffness, patients are familiar with the concept and used to completing it. Additionally, although the concept of duration was suggested to be difficult for patients, the timing of stiffness was important to patients (Chapter 4). Furthermore, the retained duration item ('Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?') was the only traditional duration item included in this study that did not focus specifically on MS. Although the item asks about stiffness on waking, the response options allow patients to describe any experience in the following 24 hour period which was consistent with the patient experience of stiffness (Chapter 4). Despite the poor performance of some traditional items, others performed well as individual items. Like many of the RAST items, the wording of 'Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week' reflects the patient suggestion that stiffness is not only relevant in the morning period (Chapters 4 and 5). Although it was not included in the RAST, 'Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week' demonstrated appropriate relationships with other measures of disease which were comparable with those reported for the RAST severity component. This provides support of the use of this item within the work into flare and within the currently unvalidated PFQ (Bykerk *et al*, 2012; Bykerk *et al*, 2014b).

An unexpected finding in this study were the low correlations between self-reported flare and all stiffness items. However, given the strong correlations demonstrated between stiffness items and disease activity, poor correlations in relation to flare may be a result of the question used to explore this relationship. The flare item used in the questionnaire pack was taken from the PFQ (Bykerk *et al*, 2012; Bykerk *et al*, 2014b).

However, the PFQ is as yet unvalidated and work into the definition and assessment of flare is ongoing, and while a core domain set for flare assessment has been endorsed at OMERACT (Bykerk *et al*, 2014a), the most appropriate way to assess these domains has not yet been established (Bartlett *et al*, 2015). Importantly, one of the RA flare core set domains is stiffness, and is one of the domains which has been identified as requiring work into most appropriate assessment (Bingham *et al*, 2011; Bykerk *et al*, 2014a). Different definitions and assessments of flare have been used in other work (Bingham et al, 2009; Lie *et al*, 2014). In further development and validation of RAST it would be relevant to consider the relationship with flare using different flare assessments. This may include the complete PFQ once it is validated. Additionally, as stiffness featured in the flare score set (Bykerk *et al*, 2014a), the use of RAST may be relevant in the context of flare assessment.

This discussion has reviewed aspects relevant to the development of the final structure and content of the RAST. Importantly this highlighted that the RAST captures the patient perspective of stiffness which is essential in PRO's. It has also identified a number of key areas for further validation and development of the RAST. Firstly, as preliminary validation was performed in the population in which it was developed, validity testing must be repeated in a new set of RA patients. This may provide an opportunity to further explore aspects of validity such as the relationship between the RAST and disease activity (e.g. the DAS28, blood tests, and flare). It would also enable exploration of the performance of RAST in relation to other psychometric properties including test-retest reliability, ability to detect change, floor and ceiling effects, and other quality criteria for evaluating questionnaires (Terwee *et al,* 2007).

9.6 Conclusion

This study has described the development of the 21-item, 3-component RAST. The process of development was rigorous and involved careful investigation from both statistical and conceptual perspectives. RAST demonstrated appropriate relationships with other measures of disease which were as good as or better than traditional stiffness items. However, the novel characteristic of RAST is its superior to face and content validity. The new RAST appears suitable for use in the assessment of stiffness and for further development and validation. Chapter 10 will now provide a summary of the key findings, strengths and limitations and implications of this body of research.

Chapter 10: Summary and discussion

This chapter summarises and draws together the main findings from the studies presented earlier and discusses the implications of this work.

10.1 Thesis aims

Stiffness is a key patient symptom for people with RA and is regularly used as an outcome measure in clinical and research settings, but currently its patient-reported assessment is variable, non-standardised, and has not been developed according to current guidelines (e.g. USDHHS FDA, 2009). The overall aim for this thesis was to explore the experience of stiffness in people with RA and use this to develop and test a new RA stiffness PROM. PROM provide an assessment of a patient's health condition that comes directly from a patient and as such are useful in the assessment of concepts that are best understood by patients, such as disease symptoms (USDHHS FDA, 2009). The justification for the development of a new stiffness PROM was the provision of a standardised assessment that captures this patient relevant symptom, and could be used in clinical and research environments. To achieve this a mixed methods approach was employed which first aimed to qualitatively understand the experience of stiffness in people with RA and then explore which aspects might be relevant in the patient-reported assessment of stiffness. Further qualitative work developed a set of items that captured those patient relevant aspects using appropriate wording and formatting and ensured the acceptability of the draft items to the target population. Using these items a quantitative survey was undertaken to provide data on which to perform multivariate analysis to develop the smallest and most internally consistent set of items to form an RA stiffness PROM and then to test how these items perform compared to current stiffness assessments. From these results it was possible to make recommendations about the most appropriate way to assess stiffness in clinical and research environments.

10.2 Contributions to knowledge

This thesis contributed to knowledge in the following ways:

- A systematic literature review of currently available stiffness assessment tools and their measurement properties identified that current stiffness assessment is based on non-standardised and poorly defined items that do not appear to have been developed in accordance with PROM development guidelines
- Investigation into the experience of stiffness for people with RA confirmed that stiffness was a significant patient symptom but also enhanced understanding

of stiffness from the patient perspective and enabled the development of a conceptual model of stiffness in RA

- Rigorous item development and testing resulted in a novel stiffness questionnaire (RAST) which was acceptable to patients, demonstrated validity during preliminary testing, and is likely to be useful in future assessment of RA stiffness in a research context
- Testing of two multivariate analytical methods (PCA and NLPCA) provided novel evidence that the two approaches produced similar results when analysing the same dataset

10.2.1 Current stiffness assessment tools and measurement properties

The systematic literature review (Chapter 2) found 19 articles all assessing stiffness from only two concepts (duration or severity), yet 37 different stiffness assessment tools were identified. This highlighted the need for standardisation of stiffness assessment in RA. On review of the measurement properties of the available tools, there was no clear evidence regarding the most appropriate measurement tool to use to assess stiffness. Additionally, no identified tools appeared to have been developed in accordance with current standards (USDHHS FDA, 2009), and there was no evidence of an appropriate conceptual framework for stiffness in RA on which to constitute stiffness assessment. Therefore the need for an RA stiffness measure with appropriate content validity was highlighted.

10.2.2 Understanding stiffness

A better understanding of the patient experience of stiffness is important because it has been a poorly understood and under-researched topic and because it is crucial for PROM development. Prior to Study 1, little was known about the patient experience of stiffness and only one previous study on the topic had been performed (Lineker *et al,* 1999), which focused on the development of a patient-centred definition of MS. Study 1 investigated the patient experience of stiffness using semi-structured interviews, and from it a conceptual model of the patient experience of stiffness was developed. This model was reinforced by data generated in Study 2 involving a different sample of patients and a different data collection method (focus groups). A qualitative study in a US-population (Orbai *et al,* 2014) was performed at a similar time to Study 1 which reported similar results (presented orally at OMERACT 12 and in a collaborative publication (Orbai *et al,* 2015)). The key similarities were that stiffness was variable within and between participants, and with respect to location

and disease activity, and was particularly important in flare. Both studies reported that participants did not experience stiffness exclusively during the morning period, and highlighted similar factors that exacerbated or alleviated stiffness. Importantly both stated that stiffness was described in terms of its impact on patients' lives. This comparison recognised important considerations in relation to stiffness assessment including investigation into cross-study aspects (impact, severity, timing, location and duration) and how these might fit into measurement, and identification of the most appropriate way to assess stiffness (Orbai *et al,* 2015). PROM development guidelines highlight the importance of underpinning qualitative studies such as these to identify relevant concepts and inform item development (USDHHS FDA, 2009). The conceptual model that emerged from Chapters 4 and 5 was supported by the similarities between the results from Study 1 and the study by Orbai *et al.* (2014) and provided a basis for PROM development. Furthermore, it identified inconsistencies between the patient experience of stiffness and current assessment (e.g. focus on EMS and limited to duration and severity).

It is helpful to consider the RA patient experience of stiffness within a broader rheumatology context, and recently published work in PMR has explored this further. PMR is a condition in which stiffness is a central symptom. Qualitative work involving eight focus groups (Mackie *et al,* 2015) previously reported in an abstract only (Hughes *et al,* 2012), developed a conceptual model of stiffness in PMR where stiffness was integral to PMR, linked to function and pain, and often discussed by patients in relation to how it impacted on their daily lives (Mackie *et al,* 2015). Twohig *et al.* (2015a) investigated the patient experience of PMR and identified 'pain, stiffness, and weakness' as one resulting theme. The similarities between the RA and PMR patient experience of stiffness were accentuated in a comment by Twohig *et al.* (2015b) published in response to the article capturing the Study 1 results (Halls *et al,* 2015). This highlighted cross-condition similarities with regards to the relationship between stiffness and pain, the conflict between the biomedical understanding and patient descriptions of the timing of stiffness, and the relevance of impact to patients (Twohig *et al,* 2015b). The adequacy of MS duration as a stiffness measure was challenged in PMR (Twohig *et al,* 2015a; Mackie *et al,* 2015) and RA (Orbai *et al,* 2014; Halls *et al,* 2015; Studies 1 and 2). Variability in PMR MS duration was reported and patients did not suggest the use of MS duration as a method of stiffness assessment (Mackie *et al,* 2015). Furthermore, the pattern of stiffness described by PMR patients was often much broader than the traditionally accepted concept of 'MS' and it was suggested that 'stiffness' may be more relevant than 'MS' (Twohig *et al,*

2015a). This was reinforced in a recent Delphi study in the development of a core domain set for PMR where patients articulated a preference for 'stiffness' rather than 'MS' (Helliwell *et al,* 2015). This is also consistent with the finding that the RA patient experience is not limited to the morning (Studies 1 and 2; Orbai *et al,* 2014; Halls *et al,* 2015). The use of the broader term 'stiffness' was highlighted as appropriate and acceptable to patients during Study 2. This is also supported by older RA literature where stiffness was found to present in the morning but also after a period of immobility (Hazes, Hayton, and Silman, 1993) and where 'immobility' was included in a patient-generated definition of 'MS' (Lineker *et al,* 1999).

The suggestion that stiffness measurement could be based on the concept of impact (Halls *et al,* 2015; Study 1) was reinforced in the work by Orbai *et al.* (2014) as impact was an area of common language in an otherwise varied symptom. It was also supported by both PMR studies (Twohig *et al,* 2015a; Mackie *et al,* 2015).

Given the similarities in the experience of stiffness and apparent shared concepts between patients with PMR and RA, there would be value in exploring the possibility of shared measurement. The specific or general nature of stiffness assessment within rheumatology is an important area for further research. The OMERACT stiffness special interest group aims to enable investigation of stiffness assessment across conditions (Orbai *et al,* 2015). Given the criticisms of traditional stiffness assessment in RA and PMR populations, an area for future research would be exploration into stiffness assessment in other rheumatic populations.

10.2.3 Development and content of a novel stiffness PROM (RAST)

The combination of qualitative (Studies 1, 2 and 3) and quantitative (Study 4) methods in the development and subsequent testing of the RAST is a key strength. The development process was congruent with recommended PROM development methodology (USDHHS FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b). The inclusion of the patient perspective throughout the studies is consistent with practices of groups such as OMERACT who advocate and implement patient involvement in outcome assessment in rheumatology (Hewlett *et al,* 2006; de Witt *et al,* 2011). It also fits with recent work highlighting the importance of patient participation in PROM development (de Wit, Kvien and Gossec, 2015). Specifically, the initial qualitative work (Studies 1 and 2) developed understanding of stiffness from the patient perspective. These data informed the iterative item development process (Chapter 6) which involved the perspectives of clinicians, patients and researchers and

considered current literature. Item content, wording and format were checked with patients during Study 3. Subsequent item reduction involved consideration of statistical criteria, theoretical appropriateness and simplicity (Nunnally and Bernstein, 1994; Pett, Lackey and Sullivan, 2003) to identify the smallest and most effective item structure (Study 4). The large VAF (71.19%) reported in the initial 38-item model (Chapter 8) emphasised the strength of the items prior to item refinement (Chapter 9) where a 21-item, three component model was defined and tested.

The PROM development process, embedded within the patient perspective, enabled the content of the RAST to reflect the patient experience of stiffness. The final RAST contained three individual components capturing stiffness severity, physical impact and psychosocial impact. The severity component contained items relating to the timing and location of stiffness, stiffness after immobility, stiffness duration and severity, and the broad impact of stiffness. The two specific impact components contained items capturing physical and daily life impact and psychosocial impact of stiffness. The content of RAST reflects the concepts identified by patients in Studies 1 and 2 and within the literature (Orbai *et al,* 2014; Orbai *et al,* 2015; Twohig *et al,* 2015a; Twohig *et al,* 2015b; Mackie *et al,* 2015). Also in relation to content, there is overlap between RAST and traditional stiffness assessment as RAST includes two traditional stiffness assessment items; one capturing duration ('Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?') and one capturing MS severity ('How would you describe the overall level of morning stiffness you have had from the time you wake up?'). This was particularly interesting given that the concept of duration had been challenged by patients (Studies 1 and 2). Both traditional items demonstrated acceptable statistical performance (Study 4) and as suggested previously, despite patients' difficulties in reporting duration, its inclusion in the final RAST may be due to familiarity with the concept or the option provided to report 'no stiffness'. The systematic literature review (Chapter 2) reported different relationships between traditional stiffness items which varied across and within concepts (severity and duration) and also across items using different wording or timeframes. In contrast, Study 4 revealed strong correlations between severity items and moderate or strong correlations between duration items, while correlations across concepts were weaker. Overall, correlations between all traditional stiffness items (apart from 'How would you describe the overall level of morning stiffness you have had from the time you wake up?' vs 'How long does your morning stiffness last from waking until maximum improvement occurs?' (3-option ordinal response), r_s=0.497) were moderate or strong regardless of concept or item format. Given the consistency

in these results, it may be that the varied relationships between items in the systematic literature review were a result of differences in wording, format and timeframe. This strengthens the argument for stiffness assessment standardisation. In another recent study, Boers *et al.* (2015) reported that MS duration and severity items correlated moderately (r=0.50, p<0.001). These results were comparable to Study 4 where severity and duration items correlated between r_s=0.497 and r_s=0.718. The authors suggested the different concepts of severity and duration, capture different aspects of RA (Boers *et al,* 2015). This suggestion is supported by Studies 1 and 2 where duration and severity were both identified within the conceptual model, and provides support for the inclusion of 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?' and 'How would you describe the overall level of morning stiffness you have had from the time you wake up?' within RAST. It may also explain the differences found between concepts in previous literature (Vliet Vlieland *et al,* 1997). This work also reported that the severity of MS demonstrated less variability than the duration of MS over 12-weeks (Vliet Vlieland *et al,* 1997), which may provide some explanation for the importance of a shorter timeframe for 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?' Given the variability in MS duration, assessment over a shorter timeframe (e.g. today) is likely to enhance accuracy. This was also reported in work in PMR (Mackie *et al,* 2015) where given the variability in MS duration reported by participants, it was suggested that this fluctuation was a possible reason for poor performance in measurement as seen in other PMR literature (e.g. Matteson *et al,* 2012). This provides further justification for the shorter timeframe for 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?' compared to other items included in the RAST. However, it would be worthwhile to investigate the different timeframes in RAST items and establish the most relevant timeframe for 'How would you describe the overall level of morning stiffness you have had from the time you wake up?'

It is important to note that the RAST does not capture stiffness as a result of permanent damage to joints. Some patients in Study 1 suggested that damaged joints (mechanical process) may be perceived to be stiff, and were different in severity and persistence from stiffness as a result of disease activity (inflammatory process). As it is traditionally considered that stiffness is an indicator of inflammatory activity in RA (e.g. Lansbury, 1956; Hazes *et al,* 1994; Soubrier *et al,* 2006), during item development (Chapter 6) it was felt important to clarify that this was the focus of the questionnaire as part of the introduction. However, two items concerned the

relationship between stiffness and damaged joints and it was hypothesised that some items may naturally group (or factor) with those items thus potentially capturing stiffness as a result of different processes. Despite this, these items performed poorly in early investigation into item suitability and were removed from subsequent testing. It may have been that this was not relevant to enough patients to be distinguished in the survey. Although, a recent study investigating radiological damage and disease-related variables reported that MS duration was associated with radiological damage (Celepkolu *et al,* 2015), these results are not consistently demonstrated in other research (van Nies *et al,* 2015). There was no significant difference between MS duration in patients with high (Larsen score <28, n=32) and low (Larsen score ≥28, n=58) erosion scores (high Larsen scores indicate more damage), and highlighted that the mechanism for the relationship is uncertain (Celepkolu *et al,* 2015). Future research could investigate whether there are differences in RAST responses between patients with different amounts of radiological damage.

10.2.4 Measurement property evidence for stiffness assessment tools

Correlations between all stiffness items (RAST and traditional stiffness items) and other measures of disease were higher than expected (Study 4). The relationship between stiffness and disability is perhaps unsurprising when considering previous work that concluded that MS was more associated with disability than with laboratory measures such as ESR (Yazici *et al,* 2004). This may be related to the observation that self-reported measures are more closely associated with other self-reported measures than with laboratory or physician-reported measures (Pincus *et al,* 1989; Taal *et al,* 1998; Westhoff *et al,* 2008; Khan *et al,* 2009). Interestingly, in PMR it was suggested that the HAQ or MHAQ could be used for assessment of stiffness impact (Mackie *et al,* 2015). Although there was an overlap in concepts and response option format between the RAST physical component and the MHAQ (Pincus *et al,* 1983), correlations were slightly lower for other components indicating that they capture different information.

Qualitative research has consistently highlighted a relationship between the patient perspective of pain and stiffness in RA (Orbai *et al,* 2014) and PMR (Mackie *et al,* 2015; Twohig *et al,* 2015a; Twohig *et al,* 2015b). Very strong correlations between morning pain and MS severity (r=0.91, p<0.0001) were reported in a recent study (Boers *et al,* 2015), although these may have been influenced by the consistency between question timeframes and the diary-based data collection. Moderate or strong relationships between all stiffness items (RAST and traditional stiffness items) and

pain were reported in Study 4. However the relationship between RAST and pain varied between components and the shared variance suggested that different components capture different information. Furthermore, moderate or strong relationships were also reported between other measures of disease and pain indicating that this was not just the case for stiffness. It has been suggested that stiffness and pain represent different concepts and that the underlying pathophysiology of the two symptoms may differ (Boers *et al,* 2015). This suggestion is reinforced by other recent work which reported that MS severity was associated with changes in PtG assessment independent of changes in pain, indicating that pain cannot be used as a substitute assessment of stiffness (Ward, Guthrie and Alba, 2015). The close relationship between symptoms in RA has been demonstrated in other work into fatigue where it was reported that some participants found it difficult to separate symptoms (Salmon, 2015). However, this does not detract from the need to assess these symptoms. Added to which, patients in Study 2 disliked the proposal that given the correlation between stiffness and pain, these could be assessed using a combined question (Boers *et al,* 2015). It is hoped that developing a stiffness assessment tool with acceptable content validity will provide better assessment of this symptom to enable proper investigation into the relationship between these symptoms. Furthermore, in an attempt to capture stiffness yet acknowledge its close relationship with pain, RAST includes 'RA stiffness' in every stem question to enhance specificity and focus on the relevant symptom.

Study 4 also provided novel evidence of a relationship between stiffness and patient-reported disease activity. This had not been investigated in detail in previous literature and the systematic literature review only identified two studies (Westhoff *et al,* 2008; Khan *et al,* 2009) that had specifically examined this relationship. Both studies demonstrated weak correlations between composite disease activity assessment and stiffness. In contrast, Study 4 demonstrated strong correlations between all stiffness items (RAST and traditional stiffness items) and patient-reported disease activity (PDAS2, Choy *et al,* 2008). It is acknowledged that this could have been influenced by the patient-reported nature of both stiffness and disease activity (Pincus *et al,* 1989; Taal *et al,* 1998; Westhoff *et al,* 2008; Khan *et al,* 2009). However, a strength is the use of the PDAS2 without EMS (Choy *et al,* 2008; Choy *et al,* 2015; Choy and Leung, 2016) to avoid circular reasoning. More recent research (Boers *et al,* 2015) has specifically investigated the relationship between MS (duration and severity) and disease activity. The study collected information using a daily diary that asked participants to record the time they woke and if they were stiff on waking ('Yes' or

'No'). If participants responded 'Yes' they were required to indicate the severity of MS on a 100mm VAS (0=not severe at all, 100=extremely severe) and also the time that MS subsided, from which the duration of MS was calculated (the difference between the time of waking and time MS subsided). Patients also reported pain severity on a 100mm VAS (0=no pain, 100=very severe pain) and collected disease activity information using the ACR core set, DAS28, and HAQ. The study reported correlations between disease activity (DAS28, ACR20) and MS duration (r=0.28, r=0.24) and MS severity (r=0.48, r=0.45), indicating that the assessment of MS adds to what is currently captured in the current RA core set (Boers *et al,* 2015). These correlations are comparable to those reported in the earlier studies (Westhoff *et al,* 2008; Khan *et al,* 2009) but different to those reported in Study 4. It may be that the differences in correlations are a result of different assessment of disease activity (e.g. DAS28 is mostly physician-reported and includes blood tests for inflammatory markers). Therefore, further investigation into the relationship between the RAST and different measures of disease activity is an important area for future research. It is also important to consider that the stiffness assessment methods, question timeframes, and data collection methods were different across studies, making direct comparison difficult. This provides further evidence for the need to standardise stiffness assessment. Given that RAST performed as well as or better than traditional stiffness items and has superior content validity, it would be an appropriate tool for future use.

The study by Boers *et al.* (2015) also reported that MS was common in patients with low (DAS28 ≤3.2) and minimal (DAS28 <2.6) disease activity. Eighty-one participants achieved low disease activity and of these, 26% reported MS duration ≥1 hour and 37% reported MS severity >10mm. Thirty-four participants achieved minimal disease activity and of these, 18% reported MS duration ≥1 hour and 30% reported MS severity >10mm. These results are similar to results reported in previous work (Hazes, Hayton, and Silman, 1993; Khan *et al,* 2009). They call into question the traditionally accepted relationship between stiffness and disease activity as they question the value of the use of stiffness assessment purely as an indicator of inflammatory activity and make the case for its use more broadly as part of routine assessment in addition to composite scores, as recommended by Boers *et al.* (2015). Furthermore, this emphasises the importance of the use of PROMs to capture the patient experience, which is valuable information in its own right. Unlike the traditional stiffness items, RAST was developed based on qualitative research with patients, consistent with recommendations regarding the development of content validity of PROMs (USDHHS

FDA, 2009; Patrick *et al,* 2011a; Patrick *et al,* 2011b). The superior content validity of RAST compared to traditional stiffness assessment is therefore advantageous from the perspective of capturing patient relevant information regardless of its relationship with other measures of disease.

10.3 Implications for PRO methodology

The RAST was developed following PROM development guidelines (USDHHS FDA, 2009). Although initially conceived for measures used to support pharmaceutical labelling claims, these guidelines provided a rigorous framework for all PROM development and for structuring this thesis. Data generated in Studies 1, 2 and 3 have supported the emphasis the guidelines place on qualitative underpinning. The guidelines regarding the development of content validity (Patrick *et al,* 2011a, Patrick *et al,* 2011b) were particularly influential, especially for tracking the development of items (Chapter 6).

Many investigators apply PCA to multivariate analysis without consideration of the nature of the dataset to which it is being applied. The statistical literature calls attention to the potential errors related to different analytical methods, and so attention was paid to the characteristics of the items. Recommended alternatives (NLPCA and polychoric correlations) were identified and two analytical methods (PCA and NLPCA) were compared in theory and practice (using the survey data) in Chapter 8. This novel approach provided evidence demonstrating similarities between results on the dataset to be analysed here. This investigation underpinned the decision to use PCA in the development of the new RA stiffness PROM, in spite of published recommendations (e.g. Streiner and Norman, 2008). This challenges those concerned with theoretical differences to explore further the circumstances in which they have practical consequences.

One reason for the similarities in results across analyses is that, although the assumptions of parametric tests (such as distribution normality) are relevant to factor analysis (Pett, Lackey and Sullivan, 2003), both exploratory and confirmatory factor analysis may be robust in circumstances when normality is violated (Gorsuch, 1983). However, this does not explain the results from the perspective that some items in the dataset were not at an interval or ratio level of measurement. This could be investigated further in future research using IRT or Rasch analysis (Tennant and Conaghan, 2007). It would also be interesting to explore whether similar RAST

content and structure emerged using polychoric correlations (Streiner and Norman, 2008; Field, 2009). If similar results were gained from all three methods (PCA, NLPCA, and polychoric correlations), this may challenge the application of approaches considered most theoretically appropriate. However, the present dataset was generated from a questionnaire which had been rigorously developed and was accounted for very well (>70% VAF) by the analysis model. It may be that theoretical differences of greater practical consequence are present in the analysis of datasets where more 'noise' is present. This could be tested in future work by adding 'noise' (e.g. random samples of data) to the current dataset.

An additional methodological implication of this work is related to the accessibility of polychoric correlations. As options to run polychoric correlations are not included in many standard software packages (and despite being accessible to download e.g. POLYMAT-C, Lorenzo-Seva and Ferrando, 2014), this process has been described as complicated (Gaskin and Happell, 2014). Therefore, additional accessibility may improve implementation of such approaches.

10.4 Implications for research

Stiffness is commonly (Kalyoncu *et al,* 2009), although decreasingly (Labitigan *et al,* 2010) measured in rheumatology research. One new area of use is research into timed-release (delayed-release) glucocorticoids, as in the CAPRA-1 and CAPRA-2 trials (Buttgereit *et al,* 2008; Buttgereit *et al,* 2013). These RCTs have demonstrated reductions in MS severity and duration following a course of TRT prednisone specifically designed for this purpose.

Further research on the CAPRA datasets has also recently been published investigating MS from the perspective of its relationship with disease activity (Boers *et al,* 2015), response following change in treatment (Alten *et al,* 2015), and improvement thresholds (Buttgereit *et al,* 2015). The study by Boers *et al.* (2015) has been described previously (10.2.3 and 10.2.4). The study by Alten *et al.* (2015) investigated patients who demonstrated no improvement in MS while taking immediate-release (IR) prednisone (in CAPRA-1), and were switched to delayed-release (DR) prednisone. The authors reported significant reductions in MS at 3, 6, and 9 months and stated that responses were comparable to patients who had continued on DR during the original study (Buttgereit *et al,* 2008). It was concluded that DR prednisone may be appropriate for use in patients who continue to experience

MS while taking IR prednisone (Alten *et al,* 2015). Buttgereit *et al.* (2015) investigated the reduction in MS duration in patients receiving IR prednisone compared to DR prednisone. The authors reported significantly higher numbers of MS reductions at all thresholds (25%, 50% and 75%) in patients receiving DR compared to IR prednisone. It was also suggested that the defined thresholds could be useful in future work regarding treatment effectiveness (Buttgereit *et al,* 2015). MS duration has also been used recently as an outcome in research into DMARD initiation. Here, MS continued to be reported in 69.2% of patients despite DMARD initiation (Strand *et al,* 2015). Although it should be remembered that MS (≥45 minutes) was an inclusion criterion for the above trials, they indicate that research using stiffness as a primary outcome measure is continuing to be performed and the results inform the development of treatment in clinical practice. MS duration remains commonly employed in these recent studies, which is problematic given the poor performance of MS duration items ('How long does your morning stiffness last from waking until maximum improvement occurs?' on a 3-option ordinal scale or in minutes and hours) in Study 4. Poor performance of such traditional items may result in inability to demonstrate treatment effects in clinical trials. Although the remaining traditional stiffness items and RAST demonstrated appropriate relationships with other measures of disease, RAST is the only PROM with content validity. Therefore it would be important to investigate whether similar results are demonstrated in trials using stiffness assessment that captures the patient experience of stiffness. RAST may be appropriate for such a task given its appropriate validity and that it contains assessment of MS duration, enabling comparison. Therefore, use in clinical trials may be a key area of future use of the new RAST. Additionally, although considerable further testing would be required, the measurement property evidence for all stiffness items (apart from 'How long does your morning stiffness last from waking until maximum improvement occurs?' on a 3-option ordinal scale or in minutes and hours) could lead to reconsideration of the decision to exclude stiffness from the ACR core set from which it was omitted (Felson *et al,* 1993). This is supported by a statement by Yazici *et al.* (2004) that study inclusion criteria often include MS yet not HAQ or pain, which unlike stiffness are both included within the ACR core set (Felson *et al,* 1993).

A final implication for research relates to the work on the assessment of flare. Although it has been proposed that stiffness should be included as a core domain to assess RA flare (Bykerk *et al,* 2014a), Study 4 demonstrated an unexpectedly poor relationship between RAST and flare. As discussed (Chapter 9), this may have been due to the flare assessment method used, or the study population which generally

represented patients at the lower end of the disease activity spectrum (Chapter 8). Given the considered relevance of stiffness in relation to flare, this is an area requiring further research. This should include testing the PFQ (Bykerk *et al,* 2012; Bykerk *et al,* 2014b) or other flare assessments in the same population of patients to provide further validation evidence for the new RAST and also to further explore the relationship between stiffness and flare.

10.5 Implications for clinical practice

A key implication for clinical practice relates to health professional and patient discourse regarding stiffness. This work has highlighted that stiffness is a relevant patient symptom, is an important part of the patient experience of RA, and has a significant impact on daily life. This reinforces the need for health professional awareness of this symptom and the importance of its recognition during clinical consultations. It was apparent that stiffness was particularly problematic for some patients, which is consistent with work in PMR where some patients reported that stiffness was the dominant symptom (Twohig *et al,* 2015a). Therefore it is important that health professionals have an awareness of the relevance of stiffness to individual patients. This work also emphasises that discussions should be worded using 'stiffness' rather than 'MS' or 'EMS'. 'Stiffness' is acceptable and relevant to patients, both in RA (Orbai *et al,* 2014; Halls *et al,* 2015) and PMR (Twohig *et al,* 2015a; Mackie *et al,* 2015).

Given completion times for the 20-item HAQ (White, Wilson and Keysor, 2011), it is estimated that the RAST will also take <10 minutes to complete. Despite attempting to identify the smallest set of items for feasibility (Boers *et al,* 1998), the 21-items may still have implications for the ease of implementation of RAST in clinical practice settings from a time perspective. Further research investigating a shorter version of RAST for use in clinical practice would be an important development of the tool. This may be facilitated by qualitative research with health professionals to help shape an improved RAST format specifically suitable for application in clinical practice. Additionally, the identification of the best performing traditional items (Study 4) provides evidence for items that could be used in this context. On the other hand, the consequence of a shorter tool would be reduced precision, especially when being used with individual (rather than groups) of patients. Therefore, this tradeoff would need to be tested as part of future work.

Additionally, the use of stiffness for diagnostic purposes has been re-evaluated in recent research by van Nies *et al.* (2015). That study included data from large European cohort studies and investigated the diagnostic value of MS in 5202 patients with arthralgia and early arthritis. In patients with arthralgia, MS duration ≥60 minutes was found to be associated with the presence of arthritis (OR 1.49 (95% CI 1.001-2.20)-2.21 (95% CI 1.33-3.69)), but the discriminatory ability was low (AUC=0.52-0.57). In patients with early arthritis, MS was associated with RA independent of other variables such as SJC (OR 1.68-1.72 (95% CI 1.03-2.74), AUC=0.64-0.68). MS duration ≥30 minutes was reported to have the optimal, although only moderate, discriminatory ability (sensitivity=74-77%, specificity=48-52%) for RA. The study concluded that in clinical practice, stiffness is useful for diagnostic purposes given its moderate discriminative ability. The acceptable performance of the MS duration item contrasts with the results from Study 4 which demonstrated poor performance of traditional duration items ('How long does your morning stiffness last from waking until maximum improvement occurs?' on a 3-option ordinal scale or in minutes and hours). However, the question used to evaluate MS duration provided an initial option for participants to report 'no stiffness' (similar to 'Were your joints stiff when you woke up today? If yes, how long did this extra stiffness last?'), which may explain the difference in the performance of this item. Furthermore, providing patients with the option to report 'no stiffness' would seem an important implication during clinical questioning. In addition, the study recognised that the use of traditional stiffness assessment using severity and duration does not appropriately capture the patient experience (Orbai *et al,* 2014). Yet it was highlighted that a tool that did capture the patient experience was not currently available and suggested that if it were, it may improve the performance of stiffness for diagnostic purposes (van Nies *et al,* 2015), therefore emphasising an important potential area of application of the new RAST.

The clinical relevance of stiffness was also emphasised in recently published recommendations regarding stiffness in Asian RA patients (Mok *et al,* 2015). These suggested routine clinical assessment of stiffness, pain and function to ensure accurate patient assessment, yet focused on 'MS' and did not provide any suggestions regarding assessment in clinical situations. The recommendations were developed specifically for an Asian population but were based on a systematic literature review including international publications, and expert opinion which predominantly involved rheumatologists based in Asia but did included a European contributor. Therefore, it is likely that they would be relevant to other populations. However, whether RAST would be appropriate for use in populations external to the

population in which it was developed would require work on translation, including investigation into whether items are conceptually equivalent across cultures (Streiner and Norman, 2008).

10.6 Future research

10.6.1 Further development and validation of RAST

Preliminary validity evidence for the RAST has been demonstrated, but further research using different measures of disease activity would enhance this. This would include comparison between RAST and DAS82 (van der Heijde *et al,* 1990), and also blood tests for common inflammatory markers such as CRP.

Although Study 4 provided some measurement property evidence for RAST, it is necessary to further this in future work. In particular it is vital to investigate the measurement properties of test-retest reliability and ability to detect change (Streiner and Norman, 2008). These are key components that require evidence in FDA guidelines (USDHHS FDA, 2009). Test-retest reliability would involve the administration of the questionnaire at two separate time points (Streiner and Norman, 2008) to test the stability of questionnaire responses over a period where the target of measurement is not expected to have changed (USDHHS FDA, 2009). Ability to detect change would test whether the questionnaire can detect changes where the target of measurement has changed (USDHHS FDA, 2009). This could be explored by using RAST in a drug intervention study such as the CAPRA trials (Buttgereit *et al,* 2008; Buttgereit *et al,* 2013). Further development in relation to measurement properties could also investigate floor and ceiling effects of RAST (whether highest or lowest scores can be achieved, Terwee *et al,* 2007) and interpretability (can qualitative meaning be interpreted from scores, SACMOT, 2002). These investigations should be performed in new samples of patients, providing further evidence in populations who were not involved in the development of the scale. Although it is important to remember that the preliminary validity testing data, though derived from the same population, did not contribute to the development of the RAST.

As addressed in Chapter 9 further research is required to develop the protocol for the use of RAST. Regarding scoring, it would be relevant to investigate the performance of the individual components in more detail. For example, are there differences in the characteristics of patients who score high or low on different components and do the components respond differently to change? Regarding the treatment of missing data

it would be important to test different approaches such as data imputation methods. This would provide evidence to enable informed decision making about recommendations regarding acceptable levels and appropriate treatment of missing data when using RAST.

Further testing using IRT would also be worthwhile. The RAST development was based on CTT which is well used and accepted in scale development (Streiner and Norman, 2008). IRT is often recommended for use within the PROM development literature (Patrick *et al,* 2011b) as it overcomes the limitations of CTT (Streiner and Norman, 2008). However, CTT and IRT can be considered complementary, therefore further investigation of RAST using IRT or Rasch analysis (Tennant and Conaghan, 2007) would enhance our understanding of the structure of RAST, and if necessary could be used to explore the response categories, and to transform ordinal level data into linear level data (Tennant and Conaghan, 2007).

10.6.2 Specific or general nature of stiffness assessment

Given the similarities in the patient experience of stiffness in RA and PMR, stiffness measurement may also be relevant in other rheumatological patient populations. Initial investigation into stiffness assessment across conditions is the current focus of the OMERACT stiffness special interest group, informed by this thesis. Recent special interest group discussion at OMERACT 2016 suggested that stiffness was relevant across rheumatic conditions and that there would be value in universal stiffness assessment. However, this would need to reflect potential differences in the patient experience across conditions. For example, patients within the special interest group highlighted that the location of stiffness would differ for those with PMR and RA and this should be reflected in the wording of items. It was also highlighted that this would be furthered by improved understanding of stiffness pathophysiology, although this is currently another area limited by lack of appropriate outcome assessment (Halls *et al,* 2016, manuscript in preparation). Stiffness assessment is also relevant more broadly, for example stiffness is reported in healthy, older populations (Sokka *et al,* 2007). The RAST may play a part in collecting normative data, possibly for populations from different communities or different age groups.

10.7 Strengths and limitations

This overview of strengths and limitations highlights the main points discussed in detail within preceding chapters.

One of the key strengths is that this work has followed published guidelines on the development of PROMs (USDHHS FDA, 2009) particularly in relation to content validity (Patrick *et al,* 2011a; Patrick *et al,* 2011b). These have provided a rigorous and systematic framework for PROM development, highlighting the importance of an underpinning of qualitative research. This enabled the development of content validity in RAST, differentiating it from other traditional stiffness assessment approaches. Furthermore the qualitative work was performed in a rigorous and transparent way including independent analysis of data by the supervisory team and patient partners (Mays and Pope, 1995; Cohen and Crabtree, 2008), and utilising COREQ guidelines to enhance rigor and transparency (Tong, Sainsbury and Craig, 2007).

Further strengths relate to the quantitative development of RAST including the comprehensive approach to deciding upon and implementing PCA, and the substantial testing of the provisional RAST for internal consistency and robustness, including repeated testing on subsets of data to ensure it was not overly dependent on the particular dataset used for development. Another strength is the careful and detailed approach to wording, presentation and comprehension of questionnaire items, ensuring each question is likely to be understood and answers will reflect patient intentions.

There are some limitations with the work presented here. Although study samples were generally representative of an RA population in the descriptives collected, a number of aspects were not considered which might influence the generalisability of results. Culture and ethnicity were not captured, few participants had low levels of education and the ability to speak English unaided was a required inclusion criterion. Furthermore, most of the subjects in the population surveyed resided in areas of relatively low deprivation. Despite this, given the consistency of results with studies containing a more diverse sample (Orbai *et al,* 2014), it is likely that the conceptual underpinning is relevant in broader populations, and efforts to address the accessibility and readability of items for a broad population were made during item development (Chapter 6). Further development and validation of RAST would include testing it in different populations (e.g. with lower education levels or English language ability) and translation and cultural adaption (Streiner and Norman, 2008).

The content and format of the RAST was primarily directed towards and derived from patients. It may be that understanding more explicitly the perspective of practicing clinicians would have resulted in some differences that might increase the utility of the

instrument in routine clinical practice. However, the perspectives of clinicians were included as part of the supervisory team and informal feedback was gained during presentation and discussion within the local clinical department. However, this is an area that should be considered as part of future development of RAST.

Finally, full validity testing of the RAST has not been carried out. While the preliminary validity testing results are promising, further validation work will be required before the RAST can be confidently recommended.

10.8 Personal reflection

Undertaking this research project and writing this thesis has been a challenging but rewarding journey. I came to this PhD process from a non-clinical background, and with limited research experience and a basic understanding of RA. Over the past four years I have learnt a great deal and developed personally and as a researcher.

I was concerned that my non-clinical background would reduce my ability to undertake this project, especially as I had very little experience of working directly with patients. However, I learnt a considerable amount from the team's patient research partners who were very open in sharing their personal experiences and helping me understand the patient perspective of RA at a human level. It was also important that I acknowledged my personal position so that I could understand my weaknesses and work to my strengths. Although I attempted to remain neutral to this during the performance of each study, especially the qualitative work, my personal background will have influenced the research process. The development of awareness, reflection on my personal background and increasing understanding of how to take account of this during qualitative studies were important while performing this work. This was supported by the input of the patient research partners and supervisory team in all aspects of this research.

The completion of Masters-level modules in critical appraisal and qualitative methods in addition to other research training opportunities through the University have enabled me to develop skills as a researcher. One of the biggest challenges for me during this process has been presentation and communication. This was an important area for personal development especially given the importance of disseminating results and sharing knowledge in research.

The research within this thesis has been presented at local and departmental events. In addition, Study 1 was presented as an oral presentation at BSR 2014 (published abstract available in Appendix BB) and has also been published as a journal article (Halls *et al,* 2015). This work has also been presented as an invited talk at the international OMERACT conference 2014 which has led to a collaborative publication (Orbai *et al,* 2015). This collaboration is currently being furthered by the stiffness special interest group who held a discussion session at OMERACT 2016 and involved further presentation of this work (Halls *et al,* 2016, manuscript in preparation). An overview of the work within this thesis was presented as an invited talk within a session convened by the researcher at BSR 2016. The systematic literature review was presented orally and as a poster at OMERACT 2016 (abstract available in Appendix CC) and Study 4 has been published as abstracts at EULAR 2016 (Appendix DD and EE). I hope to continue to present and publish other work within this thesis and to continue to develop my skills as a researcher.

10.9 Thesis summary

Stiffness is commonly experienced by people with RA and is relevant in both clinical and research contexts. Despite this, a systematic review identified that there is no clear evidence regarding the most appropriate way to assess stiffness in RA. Current stiffness assessment is not standardised, and often involves the use of unvalidated and poorly defined items, none of which appear to have been developed according to current standards including collaboration with patients. This project developed a new assessment approach based on the patient experience of stiffness. The content and structure of the new RAST was developed during a series of qualitative and quantitative studies involving people with RA.

Preliminary validity testing supports RAST as an appropriate tool to use to assess stiffness. The development of RAST is an early step in recognising stiffness as a significant, recordable patient symptom, and is also an important step forward towards standardised assessment. Further development and validation work is now required to improve evidence of the measurement properties of RAST.

References

Abramson, J.H. (1967) On the diagnostic criteria of active rheumatoid arthritis. *Journal of Chronic Diseases*. 20, pp. 275-290.

Adams, J., Chapman, J., Bradley, S. and Ryan. S.J. (2013) Literacy levels required to complete routinely used patient-reported outcome measures in rheumatology. *Rheumatology*. 52, pp. 460-464.

Ahmad, N., Ellins, J., Krelle, H. and Lawrie, M. (2014) Person-centred care: from ideas to action bringing together the evidence on shared decision making and self-management support. London: The Health Foundation. Available from: http://www.health.org.uk/publication/person-centred-care-ideas-action [Accessed 19 October 2015].

Akad, K., Solmaz, D., Sari, I., Onen, F., Akkoc, N. and Akar, S. (2013) Performance of response scales of activity and functional measures of ankylosing spondylitis: numerical rating scale versus visual analog scale. *Rheumatology International*. 33, pp. 2617-2623.

Aletaha, D., Neogi, T., Silman, A.J., Funovits, J., Felson, D.T., Bingham, C.O., Birnbaum, N.S., Burmester, G.R., Bykerk, V.P., Cohen, M.D., Ferraccioli, G., Hazes, J.M.W., Hobbs, K., Huizinga, T.W.J., Kavanaugh, A., Kay, J., Kvien, T.K., Laing, T., Mease, P., Menard, 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. and Hawker, G. (2010) 2010 Rheumatoid Arthritis Classification Criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis and Rheumatism*. 62 (9), pp. 2569-2581.

Alten, R., Holt, R., Grahn, A., Rice, P., Kent, J., Buttgereit, F. and Gibofsky, A. (2015) Morning stiffness with delayed-release prednisone after ineffective course of immediate-release prednisone. *Scandinavian Journal of Rheumatology*. 44 (5), pp. 354-358.

American College of Rheumatology. (2002) Guidelines for the management of rheumatoid arthritis: 2002 update. *Arthritis and Rheumatism*. 46 (2), pp. 328-346.

Anderson, D.L. (2001) Development of an instrument to measure pain in rheumatoid arthritis: rheumatoid arthritis pain scale (RAPS). *Arthritis Care and Research*. 45, 317-323.

Anderson, J., Caplan, L., Yazdany, J., Robbins, M.L., Neogi, T., Michaud, K., Saag, K.G., O'Dell, J.R. and Kazi, S. (2012) Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. *Arthritis Care and Research*. 64 (5), pp. 640-647.

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., Medsger, T.A., Mitchell, D.M., Neustadt, D.H., Pinals, R.S., Schaller, J.G., Sharp, J.T., Wilder, R.L. and Hunder, G.G. (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis and Rheumatism*. 31 (3), pp. 315-324.

Arthritis Research Campaign. (2005) *Collected reports on the rheumatic diseases 2005 series 4 (revised): Polymyalgia rheumatica and giant cell arteritis.* Derbyshire: Arthritis Research Campaign.

Arthritis Research UK. (2012) *Complementary and alternative medicines for the treatment of rheumatoid arthritis, osteoarthritis and fibromyalgia.* Derbyshire: Arthritis Research UK.

Arthritis Research UK. (2013a) *Practitioner-based complementary and alternative therapies for the treatment of rheumatoid arthritis, osteoarthritis, fibromyalgia and low back pain.* Derbyshire: Arthritis Research UK.

Arthritis Research UK. (2013b) *Osteoarthritis in general practice. Data and perspective.* Chesterfield: Arthritis Research UK.

Arthritis Research UK. (2014) *Sjögren's syndrome*. Chesterfield: Arthritis Research UK.

Arthur, V. and Hill, J. (2006) The musculoskeletal system and rheumatic diseases. In: Hill, A. ed. (2006) *Rheumatology nursing. A creative approach.* Chichester: John Wiley & Sons Ltd, pp. 25-92.

Arthur, S. and Nazroo, J. (2003) Designing fieldwork strategies and materials. In: Ritchie, J. and Lewis, J. eds. (2003) *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: SAGE Publications Ltd, pp. 109-137.

Arvidson, N.G., Gudbjornsson, B., Elfman, L., Ryden, A-C., Totterman, T.H. and Hallgren, R. (1994) Circadian rhythm of serum interleukin-6 in rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 53, pp. 521-524.

Arvidson, N.G., Gudbjornsson, B., Larsson, A. and Hallgren, R. (1997) The timing of glucocorticoid administration in rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 56 (1), pp. 27-31.

Attride-Stirling, J. (2001) Thematic networks: an analytic tool for qualitative research. *Qualitative Research*. 1, pp. 385-405.

Baglin, J. (2014) Improving your exploratory factor analysis for ordinal data: a demonstration using FACTOR. *Practical Assessment, Research and Evaluation* [online]. Available from: http://pareonline.net/getvn.asp?v=19&n=5 [Accessed 17 June 2015].

Bakker, M.F., Jacobs, J.W.G., Kruize, A.A., van der Veen, M.J., van Booma-Frankfort, C., Vreugdenhil, S.A., Bijlsma, J.W.J., Lafeber, F.P.J.G. and Welsing, P.M.J. (2011) Misclassification of disease activity when assessing individual patients with early rheumatoid arthritis using disease activity indices that do not include joints of feet. *Annals of the Rheumatic Diseases*. 71, pp. 830-835.

Barbour, R. (2008) *Doing Focus Groups*. London: SAGE Publications Ltd.

Barlow, J. (2001) How to use education as an intervention in osteoarthritis. *Best Practice and Clinical Rheumatology*. 15 (4), pp. 545-588.

Barlow, J., Turner, A.P. and Wright, C. (1998) Comparison of clinical and self-reported diagnoses for participants on a community-based arthritis self-management programme. *British Journal of Rheumatology*. 37, pp. 985-987.

Barlow, J., Wright, C., Sheasby, J., Turner, A. and Hainsworth, J. (2002) Self-management approaches for people with chronic conditions: a review. *Patient Education and Counselling*. 48, pp. 177-187.

Bartlett S.J., Bykerk, V.P., Cooksey, R., Choy, E.H., Alten, R., Christensen, R., Furst, D.E., Guillemin, F., Halls, S., Hewlett, S., Leong, A.L., Lyddiatt, A., March, L., Montie, P., Orbai, A-M., Pohl, C., Scholte-Voshaar, M., Woodworth, T.G. and Bingham, C.O. (2015) Feasibility and domain validation of rheumatoid arthritis (RA) flare core domain set: report of the OMERACT 2014 RA Flare Group plenary. *The Journal of Rheumatology*. 42 (11), pp. 2185-2189.

Bartlett, S.J., Hewlett, S., Bingham, C.O., Woodworth, T.G., Alten, R., Pohl, C., Choy, E.H., Sanderson, T., Boonen, A., Bykerk, V., Leong, A.L., Strand, V., Furst, D.E., Christensen, R. and The OMERACT RA Flare Working Group. (2012) Identifying core domains to assess flare in rheumatoid arthritis: an OMERACT international patient and provider combined Delphi consensus. *Annals of the Rheumatic Diseases*. 71 (11), pp. 1855-1860.

Beatty, P.C. and Willis, G.B. (2007) Research synthesis: the practice of cognitive interviewing. *The Public Opinion Quarterly*. 71 (2), pp. 287-311.

Bellamy, N. (2005) The WOMAC knee and hip osteoarthritis indices: development, validation, globalization and influence on the development of the AUSCAN hand osteoarthritis indices. *Clinical and Experimental Rheumatology*. 23 (Suppl 39), pp. 148-153.

Bellamy, N. and Buchanan, W.W. (1986) A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. *Clinical Rheumatology*. 5 (2), pp. 231-241.

Bellamy, N., Buchanan, W.W., Goldsmith, C.H., Campbell, J. and Stitt, L. (1988) Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *The Journal of Rheumatology*. 15, pp. 1833-1840.

Bellamy, N., Campbell, J., Haraoui, B., Buchbinder, R., Hobby, K., Roth, J.H. and MacDermid, J.C. (2002) Dimensionality and clinical importance of pain and disability in hand osteoarthritis: development of the Australian/Canadian (AUSCAN) Osteoarthritis Hand Index. *Osteoarthritis and Cartilage*. 10, pp. 855-862.

Bennett, R.M., Friend, R., Jones, K.D., Ward, R., Han, B.K. and Ross, R.L. (2009) The Revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties. *Arthritis Research and Therapy* [online]. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2745803/pdf/ar2783.pdf [Accessed 4 March 2015].

Bernstein, I.H. and Teng, G. (1989) Factoring items and factoring scales are different: spurious evidence for multidimensionality due to item categorisation. *Psychological Bulletin*. 105 (3), pp. 467-477.

Berthelot, J.M., Bernelot-Moens, H.J., Klarlund, M., McGonagle, D., Calin, A., Schumacher, H.R., Combe, B., De Bandt, M., Drosos, A.A., Flipo, R.M., Harrison, B.J., Kaarela, K., Le Goff, P., Meyer, O., Punzi, L., Zerbini, C.A., Saraux, A. and the CRI Group. (2002) Differences in understanding and application of 1987 ACR criteria for rheumatoid arthritis and 1991 ESSG criteria for spondylarthropathy. A pilot survey. *Clinical and Experimental Rheumatology*. 20, pp. 145-150.

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., Pohl, C., Sanderson, T.C., Strand, V. and 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 (8), pp. 1751-1758.

Bingham, C.O., 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. and Alten, R. (2009) Developing a standardized definition for disease “flare” in rheumatoid arthritis (OMERACT 9 special interest group). *The Journal of Rheumatology*. 36, pp. 2335-2341.

Birks, M. and Mills, J. (2011) *Grounded Theory: A Practical Guide*. London: SAGE Publications Ltd.

Bloor, M. (2001) *Focus Groups in Social Research*. London: SAGE Publications Ltd.

Boers, M., Brooks, P., Strand, C.V. and Tugwell, P. (1998) The OMERACT Filter for outcome measures in rheumatology. *The Journal of Rheumatology*. 25 (2), pp. 198-199.

Boers, M., Buttgereit, F., Saag, K., Alten, R., Grahn, A., Storey, D., Rice, P. and Kirwan, J. (2015) What Is the relationship between morning symptoms and measures of disease activity in patients with rheumatoid arthritis? *Arthritis Care and Research*. 67 (9), pp. 1202-1209.

Boers, M., Kirwan, J.R., Gossec, L., Conaghan, P.G., D'Agostino, M-A., Bingham, C.O., Brooks, P.M., Landewé, R., March, L., Simon, L., Singh, J.A., Strand, V., Wells, G.A. and Tugwell, P. (2014) How to choose core outcome measurement sets for clinical trials: OMERACT 11 approves Filter 2.0. *The Journal of Rheumatology*. 41 (5), pp. 1025-1030.

Boers, M., Kirwan, J.R., Wells, G., Beaton, D., Gossec, L., d'Agostino, M.A., Conaghan, P.G., Bingham, C.O., Brooks, P.M., Landewé, R., March, L., Simon, L., Singh, J.A., Strand, V. and Tugwell, P. (2014) Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *Journal of Clinical Epidemiology*. 67 (7), pp. 745-753.

Bollen, K.A. and Barb, K.H (1981) Pearson’s R and coarsely categorised measures. *American Sociological Review*. 46 (2), pp. 232-239.

Bombardier, C., Barbieri, M., Parthan, A., Zack, D.J., Walker, V., Macarios, D., Smolen, J.S. (2012) The relationship between joint damage and functional disability in rheumatoid arthritis: a systematic review. *Annals of the Rheumatic Diseases*. 71, pp. 836-844.

Borstlap, M., Zant, J.L., van Soesbergen, R.M. and van der Korst, J.K. (1995) Quality of life assessment a comparison of four questionnaires for measuring improvements after total hip replacement. *Clinical Rheumatology*. 14 (1), pp. 15-20.

Boyatzis, R.E. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development.* London: SAGE Publications Ltd.

Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*. 3 (2), pp. 77-101.

Braun, V. and Clarke, V. (2012) *Successful Qualitative Research*. London: SAGE Publications Ltd.

Brett, J., Staniszewska, S., Mockford, C., Seers, K., Herron-Marx, S and Bayliss, H. (2010) The PIRICOM Study: A systematic review of the conceptualisation, measurement, impact and outcomes of patients and public involvement in health and social care research. *UKCRC* [online]. Available from: http://www.ukcrc.org/wp-content/uploads/2014/03/Piricom+Review+Final+2010.pdf [Accessed 20 November 2012].

Britten, N. (2006) Qualitative interviews. In: Pope, C. and Mays, N., eds. (2006) *Qualitative Research in Health Care*. Oxford: Blackwell Publishing, pp. 12-20.

Bryman, A. (2006) Integrating quantitative and qualitative research: how is it done? *Qualitative Research*. 6 (1), pp. 97-113.

Bryman, A. (2007) Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research*. 1 (1), pp. 8-22.

Buchbinder, R., Bombardier, C., Yeung, M. and Tugwell, P. (1995) Which outcome measures should be used in rheumatoid arthritis clinical trials? Clinical and quality-of-life measures' responsiveness to treatment in a randomized controlled trial. *Arthritis and Rheumatism*. 38 (11), pp. 1568-1580.

Burckhardt, C.S. Clark, S.R. and Bennett, R.M. (1991) The fibromyalgia impact questionnaire: development and validation. *The Journal of Rheumatology*. 18 (5), pp. 728-33.

Burton, S. and Lloyd, M. (2006) An overview of rheumatoid arthritis. *Nursing Standard*. 20 (24), pp. 46-49.

Buttgereit, F. (2011) How should impaired morning function in rheumatoid arthritis be treated? *Scandinavian Journal of Rheumatology*. 40 (Suppl 125), s28-s39.

Buttgereit, F., Doering, G., Schaeffler, A., Witte, S., Sierakowski, S., Gromnica-Ihle, E., Jeka, S., Krueger, K., Szechinski, J. and Alten, R. (2008) Efficacy of modified-release versus standard prednisone to reduce duration of morning stiffness of the joints in rheumatoid arthritis (CAPRA-1): a double-blind, randomised controlled trial. *The Lancet*. 371, pp. 205-214.

Buttgereit, F., Kent, J.D., Holt, R.J., Grahn, A.Y., Rice, P., Alten, R. and Yazici, Y. (2015) Improvement thresholds for morning stiffness duration in patients receiving delayed- verses immediate-release prednisone for rheumatoid arthritis. *Bulletin of the Hospital for Joint Diseases*. 73 (3), pp. 168-177.

Buttgereit, F., Mehta, D., Kirwan, J., Szechinski, J., Boers, M., Alten, R.E., Supronik, J., Szombati, I., Romer, U., Witte, S. and Saag, K.G. (2013) Low-dose prednisone chronotherapy for rheumatoid arthritis: a randomised clinical trial (CAPRA-2). *Annals of the Rheumatic Diseases.* 72, pp. 204-210.

Bykerk, V.P., Bartlett, S.J., Choy, E., Hitchon, C., Pope, J., Thorne, C., Haraoui, B., Keystone, E. and Bingham, C.O., on behalf of the OMERACT RA flare group and CATCH. (2012) An evaluation of flare in patients with early rheumatoid arthritis using OMERACT preliminary flare questionnaire. *Annals of the Rheumatic Diseases*. 71 (Suppl 3), pp. 180-181.

Bykerk, V.P. Bingham, C.O., Choy, E.H., Boire, G., Lin, D., Pope, J.E., Thorne, J., Hitchon, C., Keystone, E.C. and Bartlett, S.J., on behalf of OMERACT Flare Group, on behalf of CATCH Investigators. (2014b) The OMERACT preliminary flare questionnaire (PFQ) is responsive to change and able to detect clinically important worsening indicating need for treatment change in the Canadian early arthritis cohort. *Annals of the Rheumatic Diseases*. 73 (Suppl 2), pp. 418-419.

Bykerk, V.P., Lie, E., Bartlett, S.J., Alten, R., Boonen, A., Christensen, R., Furst, D.E., Hewlett, S., Leong, A.L., Lyddiatt, A., Marsh, L., May, J.E., Montie, P., Orbai, A-M., Pohl, C., Scholte-Voshaar, M., Woodworth, T., Bingham, C.O., Choy, E.H. (2014a) Establishing a core domain set to measure rheumatoid arthritis flares: report of the OMERACT 11 RA flare workshop. *The Journal of Rheumatology*. 41 (4), pp. 799-809.

Calin, A., Nakache, J-P., Gueguen, A., Zeidler, H., Mielants, H. and Dougados, M. (1999) Defining disease activity in ankylosing spondylitis: is a combination of variables (Bath ankylosing spondylitis disease activity index) an appropriate instrument? *Rheumatology*. 38, pp. 878-882.

Celepkolu, T., Sariyildiz, M.A. Yazmalar, L., Bucaktepe, P.G.E., Cetincakmak, M.G. and Satici, O. (2015) Radiological damage in rheumatoid arthritis: relationship between the duration of morning stiffness, depressive symptoms, and disease duration. *Acta Medica Mediterranea*. 31, pp. 251-256.

Chambers, R. (2006) The role of the health professional in supporting self care. *Quality in Primary Care*. 14, pp. 129-131.

Chapman, J.A. and Langridge, J. (1997) Physiotherapy health education literature. *Physiotherapy*. 83 (8), pp. 406-412.

Charmaz, K. (1990) 'Discovering' chronic illness: using grounded theory. *Social Science and Medicine*. 30 (11), pp. 1161-172.

Chen, J. and Choi, J. (2009) A comparison of maximum likelihood and expected a posteriori estimation for polychoric correlation using Monte Carlo simulation. *Journal of Modern Applied Statistical Methods*. 8 (1), pp. 337-354.

Choi, J., Chen, J. and Kim, S. (2009). BayesPCC (v 3.0.0): A standalone Microsoft windows software program for estimating polychoric correlation matrices using Bayesian estimation methods [Computer software].

Choi, J., Peters, M. and Mueller, R.O. (2010) Correlational analysis of ordinal data: from Peason's *r* to Bayesian polychoric correlation. *Asia Pacific Education Review*. 11, pp. 459-466.

Choy, E. (2012) Understanding the dynamics: pathways involved in the pathogenesis of rheumatoid arthritis. *Rheumatology*. 51 (Suppl 5), pp. v3-v11.

Choy, E.H., Khoshaba, B., Cooper, D., MacGregor, A. and Scott, D.L. (2008) Development and validation of a patient-based disease activity score in RA that can be used in clinical trials and routine practice. *Arthritis Care and Research*. 59 (2), pp. 192-199.

Choy, E.H., Khoshaba, B., Cooper, D., MacGregor, A. and Scott, D.L. (2015) Missing constant intercept in the equation for calculating PDAS2 in the article by Choy et al. (*Arthritis Care and Research*, February 2008). *Arthritis Care and Research.* 67 (11), pp. 1618.

Choy, E.H. and Leung, A.M.H. (2016) Simplified patient-derived disease activity score (SPDAS2): a simplified version without early morning stiffness. *Annals of the Rheumatic Diseases*. 75 (Suppl 2), pp. 209.

Clarke, L.L., Jessop, D.S., Hunt, L.P., Straub, R.H., Perry, M.G. and Kirwan, J.R. (2011) Alleviation of morning joint stiffness by low dose prednisone in rheumatoid arthritis is associated with circadian changes in IL-6 and cortisol. *International Journal of Clinical Rheumatology*. 6, pp. 241-249.

Cleeland, C.S. and Ryan, K.M. (1994) Pain assessment: global use of the brief pain inventory. *Annals of the Academy of Medicine, Singapore*. 23 (2), pp. 129-138.

Cohen, D.J. and Crabtree, B.F. (2008) Evaluative criteria for qualitative research in health care: controversies and recommendations. *Criteria for Qualitative Research*. 6 (4), pp. 331-339.

Collins, D. (2003) Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*. 12, pp. 229-238.

Comrey, A.L. and Lee, H.B. (1992) *A First Course in Factor Analysis*. London: Psychology Press, Taylor and Francis Group.

Connolly, J., Condell, J., Curran, K. and Gardiner, P. (2012) A new method to determine joint range of movement and stiffness in rheumatoid arthritic patients. *Engineering in Medicine and Biology Society (EMBC) conference proceedings IEEE* [online]. Available from: http://ieeexplore.ieee.org/xpl/abstractAuthors.jsp?arnumber=6347455 [Accessed 11 April 2014].

Conrad, F. and Blair, J. (1996) From impressions to data: increasing the objectivity of cognitive interviews. *Proceedings of the Section on Survey Research Methods, American Statistical Association [online]. Available from:* http://www.amstat.org/sections/srms/Proceedings/ *[Accessed 7 April 2015].*

Conrad, F. and Blair, J. (2009) Sources of error in cognitive interviews. *Public Opinion Quarterly*. 73 (1), pp. 32-55.

Cook, L.L., Eignor, D.R. and Taft, H.L. (1988) A comparative study of the effects of recency of instruction on the stability of IRT and conventional item parameter estimates. *Journal of Educational Measurement*. 25 (1), pp. 31-45.

Cornell, P. (2007) Management of patients with rheumatoid arthritis. *Nursing Standard*. 22 (4), pp. 51-57.

Creswell, J.W. (2003) *Research Design. Qualitative, Quantitative and Mixed Methods Approaches*. London: SAGE Publications Ltd.

Creswell, J.W. and Clark, V.L.P. (2011) *Designing and Conducting Mixed Methods Research.* London: SAGE Publications Ltd.

Cutolo, M. (2011) How should morning function in rheumatoid arthritis be assessed? Bibliographic study of current assessment. *Scandinavian Journal of Rheumatology*. 40 (Suppl 125), pp. 17-22.

Cutolo, M., Iaccarino, L., Doria, A., Govoni, M., Sulli, A. and Marcassa, C. (2013) Efficacy of the switch to modified-release prednisone in rheumatoid arthritis patients treated with standard glucocorticoids. *Clinical and Experimental Rheumatology*. 31 (4), pp. 498-505.

Cutolo, M., Seriolo, B., Craviotto, C., Pizzorni, C. and Sulli, A. (2003). Circadian rhythms in RA. *Annals of the Rheumatic Diseases*. 62 (7), pp. 593-596.

Dalakas, M.C. (2009) Stiff person syndrome: advances in pathogenesis and therapeutic interventions. *Current Treatment Options in Neurology*. 11 (2), pp. 102-110.

Dancey, C.D. and Reidy, J. (2007) *Statistics Without the Maths for Psychology*. London: Prentice Hall.

Dasgupta, B., Cimmino, M.A., Kremers, H.M., Schmidt, W.A., Schirmer, M., Salvarani, C., Bachta, A., Dejaco, C., Duftner, C., Slott Jensen, H., Duhaut, P., Poór, G., Kaposi, P.N., Mandl, P., Balint, P.V., Schmidt, Z., Iagnocco, A., Nannini, C., Cantini, F., Macchioni, P., Pipitone, N., del Amo, M., Espígol-Frigolé, G., Cid, M.C., Martínez-Taboada, V.M., Nordborg, E., Direskeneli, H., Aydin, S.Z., Ahmed, K., Hazelman, B., Silverman, B., Pease, C., Wakefield, R.J., Luqmani, R., Abril, A., Michet, C.J., Marcus, R., Gonter, N.J., Maz, M., Carter, R.E., Crowson, C.S. and Matteson, E.L. (2012) 2012 Provisional classification criteria for polymyalgia rheumatica: a European League Against Rheumatism/American College of Rheumatology collaborative initiative. *Arthritis and Rheumatism*. 64 (4), pp. 943-954.

da Silva, J., Phillips, S. and Buttgereit, F. (2011) Impact of impaired morning function on the lives and well-being of patients with rheumatoid arthritis. *Scandinavian Journal of Rheumatology*. 40 (125), pp. 6-11.

Dawes, R.M. (1979) The robust beauty of improper linear models in decision making. *American Psychologist*. 34 (7), pp. 571-582.

de Champlain, A.F. (2010) A primer on classical test theory and item response theory for assessments in medical education. *Medical Education*. 44 (1), pp. 109-117.

Dekkers, C.J., Geenan, R., Godaert, G.L.R., van Doornen, L.J.P. and Bijlsma, J.W.J. (2000) Diurnal courses of cortisol, pain, fatigue, negative mood, and stiffness in patients with recently diagnosed rheumatoid arthritis. *International Journal of Behavioral Medicine*. 7 (4), pp. 353-371.

Department for Communities and Local Government. (2011) *The English indices of deprivation 2010*. London: Department for Communities and Local Government.

Department of Health (2010a) *Equity and excellence: liberating the NHS*. London: The Stationary Office. Available from: https://www.gov.uk/government/publications/liberating-the-nhs-white-paper [Accessed 1 December 2015].

Department of Health (2010b) *Self care – a real choice. Self care support – a practical option.* London: The Department of Health. Available from: http://webarchive.nationalarchives.gov.uk/20130107105354/http:/www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4101702.pdf [Accessed 19 October 2015].

Denscombe, M. (2008) Communities of practice: a research paradigm for the mixed methods approach. *Journal of Mixed Methods Research*. 2, pp. 270-283.

de Silva, M., Binder, A. and Hazleman, B.L. (1984) The timing of prednisolone dosage and its effect on morning stiffness in rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 43, pp. 790-793.

de Souza, S., Williams, R. and Lempp, H. (2016) Patient and clinician views on the quality of foot health care for rheumatoid arthritis outpatients: a mixed methods service evaluation. *Journal of Foot and Ankle Research* [online]. Available from: http://jfootankleres.biomedcentral.com/articles/10.1186/s13047-015-0133-2 [Accessed 19 October 2016].

DeVellis, R.F. (2012) *Scale Development. Theory and Applications*. London: SAGE Publications Ltd.

de Wit, M.P.T., Abma, T.A., Koelewijn-van Loon, M.S., Collins, S. and Kirwan, J. (2014) What has been the effect on trial outcome assessments of a decade of patient participation in OMERACT? *The Journal of Rheumatology*. 41 (1), pp. 177-184.

de Wit, M.P.T., Berlo, S.E., Aanerud, G.J., Aletaha, D., Bijlsma, J.W., Croucher, L., Da Silva, J.A.P., Glüsing, B., Gossec, L., Hewlett, S., Jongkees, M., Magnusson, D., Scholte-Voshaar, M., Richards, P., Ziegler, C., Abma, T.A. (2011) European League Against Rheumatism recommendations for the inclusion of patient representatives in scientific projects. *Annals of the Rheumatic Diseases*. 70, pp. 722-726.

de Wit, M.P.T., Kvien, T.K. and Gossec, L. (2015) Patient participation as an integral part of patient-reported outcomes development ensures the representation of the patient voice: a case study from the field of rheumatology. *Rheumatic and Musculoskeletal Diseases Open* [online]. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4613173/pdf/rmdopen-2015-000129.pdf [Accessed 11 November 2015].

Dickens, C., McGowan, L., Clark-Carter, D. and Creed, F. (2002) Depression in rheumatoid arthritis: a systematic review of the literature with meta-analysis. *Psychosomatic Medicine*. 64, pp. 52-60.

DiRaimondo, V.C. and Forsham, P.H. (1958) Pharmacophysiologic principles in the use of corticosteroids and adrenocorticotrophin. *Metabolism*. 7, pp. 5-24.

Drennan, J. (2003) Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *Methodological Issues in Nursing Research*. 42 (1), pp. 57-63.

Dures, E., Almeida, C., Caesley, J., Peterson, A., Ambler, N., Morris, M., Pollock, J. and Hewlett, S. (2014). A Survey of Psychological Support Provision for People with Inflammatory Arthritis in Secondary Care in England. *Musculoskeletal Care*. 12 (3), pp.173-181.

Dures, E.K., Hewlett, S., Cramp, F.A., Greenwood, R., Nicklin, J.K., Urban, M. and Kirwan, J.R. (2013) Reliability and sensitivity to change of the Bristol Rheumatoid Arthritis Fatigue Scales. *Rheumatology*. 52 (10), pp. 1832-1839.

Dwek, M-R., McBain, H., Cleanthous, S., Shipley, M. and Newman, S. (2015) Osteoarthritis: patient expectations about future Pain, stiffness and function. *Musculoskeletal Care*. 13 (2), pp. 84-92.

Dziedzic, K.S.G. (1997) *The body chart: a futher sketch towards a fuller picture of ankylosing spondylitis* [online]. PhD, Keele University. Available from: http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.388352 [Accessed 24 April 2014].

Elasy, T.A. and Gaddy, G. (1998). Measuring subjective outcomes. Rethinking reliability and validity. *The Journal of General Internal Medicine*. 13, pp. 757-761.

El Miedany, Y., El Gaafary, M., Youssef, S. and Palmer, D. (2010) Incorporating patient reported outcome measures in clinical practice: development and validation of a questionnaire for inflammatory arthritis. *Clinical and Experimental Rheumatology*. 28 (5), pp. 734-744.

Embretson, S.E. and Hershberger, S.L. (1999) Summary and future psychometric models in testing. In Embretson, S.E. and Hershberger, S.L. eds. (1999) *The New Rules of Measurement*. London: Lawrence Erlbaum Associates, Publishers, pp. 243-254.

Fabrigar, L.R. Wegener, D.T., MacCallum, R.C. and Strahan, E.J. (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*. 4 (3), pp. 272-299.

Fanning, E. (2005) Formatting a paper-based survey questionnaire: best practices. *Practical Assessment, Research and Evaluation* [online]. Available from: http://pareonline.net/genpare.asp?wh=0&abt=10 [Accessed 13 October 2013].

Farah, J.T., Young, J.P., LaMoreaux, L., Werth, J.L. and Poole, M. (2001) Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 94, pp. 149-58.

Felson, D.T., Anderson, J.J., Boers, M., Bombardier, C., Chernoff, M, Fried, B., Furst, D., Goldsmith, C., Kieszak, S., Lightfoot, R., Paulus, H., Tugwell, P., Weinblatt, C., Widmark, R., Williams, H.J. and Wolfe, F. (1993) The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis and Rheumatism*. 36 (6), pp. 729-740.

Felson, D.T., Smolen, J.S., Wells, G., Zhang, B., van Tuyl, L.H.D., Funovits, J., Aletaha, D., Allaart, C.F., Bathon, J., Bombardieri, S., Brooks, P., Brown, A., Matucci-Cerinic, M., Choi, H., Combe, B., de Wit, M., Dougados, M., Emery, P., Furst, D., Gomez-Reino, J., Hawker, G., Keystone, E., Khanna, D., Kirwan, J., Kvien, T.K., Landewé, R., Listing, J., Michaud, K., Martin-Mola, E., Montie, P., Pincus, T., Richards, P., Siegel, J.N., Simon, L.S., Sokka, T., Strand, V., Tugwell, P., Tyndall, A., van der Heijde, D., Verstappen, S., White, B., Wolfe, F., Zink, A. and Boers, M. (2011) American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis and Rheumatism*. 63 (3), pp. 573-586.

Field, A. (2009) *Discovering Statistics Using IBM SPSS Statistics*. London: SAGE Publications Ltd.

Finch, H. and Lewis, J. (2003) Focus groups. In Ritchie, J. And Lewis, J. eds. (2003) *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: SAGE Publications Ltd, pp. 170-198.

Finlay, L. (1998) Reflexivity: an essential component for all research? *British Journal of Occupational Therapy*. 61 (10), pp. 453-456.

Finlay, L. (2002) ''Outing'' the researcher: the provenance, process, and practice of reflexivity. *Qualitative Health Research*. 12 (4), pp. 531-545.

Fitzpatrick, R., Davey, C., Buxton, M.J. and Jones, D.R (1998) Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* [online]. Available from: http://www.journalslibrary.nihr.ac.uk/__data/assets/pdf_file/0003/64830/FullReport-hta2140.pdf [Accessed 23 November 2012].

Flurey, C. (2012) *Capturing daily fluctuations, flare and self-management in rheumatoid arthritis: The patient perspective*. PhD, University of the West of England.

Foot, C., Gilburt, H., Dunn, P., Jabbal, J., Seale, B., Goodrich, J., Buck, D. and Taylor, J. (2014) *People in control of their own health and care. The state of involvement*. London: The Kings Fund. Available from: http://www.kingsfund.org.uk/publications/people-control-their-own-health-and-care [Accessed 1 December 2015].

Fowler, F.J. (1995) *Improving Survey Questions*. London: Sage Publications Ltd.

Franchignoni, F., Salaffi, F., Ciapetti, A. and Giordano, A. (2014) Searching for optimal rating scales in the Bath Ankylosing Spondylitis Functional Index (BASFI) and Bath Ankylosing Spondylitis Disease Activity Index (BASDAI). *Rheumatology International*. 34, pp. 171-173.

Frank, G. (1997) Is there life after categories? Reflexivity in qualitative research. *Occupational Therapy Journal of Research*. 17, pp. 84-98

Fransen, J., Langenegger, T., Michel, B.A. and Stucki, G. (2000) Feasibility and validity of the RADAI, a self-administered rheumatoid arthritis disease activity index. *Rheumatology*. 39 (3), pp. 321-327.

Fransen, J., Stucki, G. and van Riel, P.L.C.M. (2003) Rheumatoid arthritis measures. *Arthritis Care and Research*. 49, (Suppl 5), pp. s214-s224.

French, T., Hewlett, S., Kirwan, J. and Sanderson, T. (2013) Different wording of the patient global visual analogue scale (PG-VAS) affects rheumatoid arthritis patients' scoring and the overall disease activity score (DAS28): a cross-sectional study. *Musculoskeletal Care*. 11, pp. 229-237.

Fries, J.F., Bruce, B., Bjorner, J. and Rose, M. (2006) More relevant, precise, and efficient items for assessment of physical function and disability: moving beyond the classic instruments. *Annals of the Rheumatic Diseases*. 65 (Suppl 3), pp. iii16-iii21.

Fries, J.F., Spitz, P., Kraines, R.G. and Holman, H.R. (1980) Measurement of patient outcome in arthritis. *Arthritis and Rheumatism*. 23 (2), pp. 137-145.

Frost, M.H., Reeve, B.B., Liepa, A.M., Stauffer, J.W., Hays, R.D. and The Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. (2007) What is sufficient evidence for the reliability and validity of patient-reported outcome measures. *Value in Health*. 10 (Suppl 2), pp. 94-105.

Gagliese, L., Weizblit, N., Ellis, W. and Chan, V.W. (2005) The measurement of postoperative pain: a comparison of intensity scales in younger and older surgical patients. *Pain*. 117 (3), pp. 412-20.

Garrett, S., Jenkinson, T., Kennedy, L.G., Whitelock, H., Gaisford, P. and Calin, A. (1994) A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. *The Journal of Rheumatology*. 21 (12), pp. 2286-2291.

Gaskin, C.J. and Happell, B. (2014) On exploratory factor analysis: a review of recent evidence, an assessment of current practice, and recommendations for future use. *International Journal of Nursing Studies*. 51, pp. 511-521.

Geenen, R., Newman, S., Bossema, E.R., Vriezekolk, J.E. and Boelen, P.A. (2012) Psychological interventions for patients with rheumatic diseases and anxiety or depression. *Best Practice and Research Clinical Rheumatology*. 26, pp. 305-319.

GeoConvert (2007) *GeoConvert* [computer programme]. Available from: http://geoconvert.mimas.ac.uk/index.html [Accessed 26 October 2015].

Glaser, B.G. and Strauss, A.L. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Publishing Company.

Goddard, R., Dowson, D., Longfield, M.D. and Wright, V. (1970) Stiffness of the knee in normal and osteoarthrosic subjects. *Annals of the Rheumatic Diseases*. 29, pp. 194.

Goodacre, L., Smith, J., Middis, D. and Goodacre, J. (2007) Development and validation of a patient-centred measure of activity limitation (MAL) in rheumatoid arthritis. *Rheumatology.* 46, pp. 703-708.

Goodwin, D. (2006) Ethical issues. In: Pope, C. and Mays, N. eds. (2006) *Qualitative Research in Health Care.* Oxford: Blackwell Publishing, pp. 53-62.

Gorsuch, R.L. (1983) *Factor analysis*. London: Lawrence Erlbaum Associates, Publishers.

Gossec, L., Dougados, M., Rincheval, N., Balanescu, A., Boumpas, D.T., Canadelo, S., Carmona, L., Daures, J-P., de Wit, M., Dijkmans, B.A.C., Englbrecht, M., Gunendi, Z., Heiberg, T., Kirwan, J.R., Mola, E.M., Matucci-Cerinic, M., Otsa, K., Schett, G., Sokka, T., Wells, G.A., Aanerud, G.J., Celano, A., Dudkin, A., Hernandez, C., Koutsogianni, K., Akca, F.N., Petre, A-M., Richards, P., Scholte-Voshaar, M., von Krause, G. and Kvien, T.K. (2009) Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative. *Annals of the Rheumatic Diseases*. 68, pp. 1680-1685.

Gossec, L., Paternotte, S., Aanerud, G.J., Balanescu, A., Boumpas, D.T., Carmona, L., de Wit, M., Dijkmans, B.A.C., Dougados, M., Englbrecht, M., Gogus, F., Heiberg, T., Hernandez, C., Kirwan, J.R., Mola, E.M., Matucci-Cerinic, M., Otsa, K., Schett, G., Scholte-Voshaar, M., Sokka, T., von Krause, G., Wells, G.A. and Kvien, T.K. (2011) Finalisation and validation of the rheumatoid arthritis impact of disease score, a patient-derived composite measure of impact of rheumatoid arthritis: a EULAR initiative. *Annals of the Rheumatic Diseases*. 70, pp. 935-942.

Green, S.B., Thompson, M.S., Levy, R. and Lo, W-J. (2015) Type I and type II error rates and overall accuracy of the revised parallel analysis method for determining the number of factors. *Educational and Psychological Measurement*. 75 (3), pp. 428-457.

Grønning, K., Rødevand, E. and Steinsbekk A. (2010) Paid work is associated with improved health-related quality of life in patients with rheumatoid arthritis. *Clinical Rheumatology*. 29, pp. 1317–1322.

Groves, R.M. Fowler, F.J., Couper, M.P., Lepowski, J.M., Singer, E. and Tourangeau, R. (2009) Survey Methodology. Chichester: John Wiley & Sons Ltd.

Guest, G., Bunce, A. and Johnson, L. (2006) How many interviews are enough?: an experiment with data saturation and variability. *Field Methods*. 18 (1), pp. 59-82.

Gunning, R. (1969) The fog index after twenty years. *Journal of Business Communication*. 6, pp. 3-13.

Hadavi, S., Noyce, A.J., Leslie, D.R. and Giovannoni, G. (2011) Stiff person syndrome. *Practical Neurology*. 11, pp. 272-282.

Haigh, R.C., McCabe, C.S., Halligan, P.W. and Blake, D.R. (2003) Joint stiffness in a phantom limb: evidence of central nervous system involvement in rheumatoid arthritis. *Rheumatology*. 42, pp. 888-892.

Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1995) *Multivariate data analysis with readings*. Englewood Cliffs: Prentice Hall.

Hakim, A., Clunie, G. and Haq, I. (2011) *Oxford Handbook of Rheumatology.* 3rd ed. Oxford University Press.

Halls, S., Dures, E., Kirwan, J., Pollock, J., Baker, G., Edmunds, A. and Hewlett, S. (2015) Stiffness is more than just duration and severity: a qualitative exploration in people with rheumatoid arthritis. *Rheumatology*. 54 (4), pp. 615-622.

Halls, S., Alten, R., Bartlett, S., Bingham, C., Campbell, I., Dures, E., Hewlett, S., Hill, C., Holt, R., Hughes, R., Kirwan, J., Leong, A., Leung, K., Lyddiatt, A., Mackie, S., March, L., Orbai, A-M., and Sinnathurai, R. (2016) Report from the OMERACT 2016 Stiffness Special Interest Group: current investigations and future research directions. *Manuscript in preparation*.

Hambleton, R.K. and Jones, R.W. (1993) An NCME instructional module on comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*. 12 (3), pp. 38-47.

Hammond, A., Jones, V. and Prior, Y. (2015) The effects of compression gloves on hand symptoms and hand function in rheumatoid arthritis and hand osteoarthritis: a systematic review. *Clinical Rehabilitation* [online]. Available from: http://cre.sagepub.com/content/early/2015/03/20/0269215515578296.long [Accessed 28 November 2015].

Hammond, A., Tennant, A., Tyson, S., Nordenskio, U., Hawkins, R. and Prior, Y. (2015) The reliability and validity of the English version of the evaluation of daily activity questionnaire for people with rheumatoid arthritis. *Rheumatology*. 54, pp. 1605-1615.

Hammond, A., Young, A. and Kidao, R. (2004) A randomised controlled trial of occupational therapy for people with early rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 63, pp. 23-30.

Happell, B. (2007) Focus groups in nursing research: an appropriate method or the latest fad? *Nurse Researcher*. 14 (2), pp. 18-24.

Harkness, J.A.L., Richter, M.B., Panayi, G.S., van de Pette, K., Unger, A., Pownall, R. and Geddawi, M. (1982) Circadian variation in disease activity in rheumatoid arthritis. *British Medical Journal*. 284, pp.551-554.

Harvey, A.R. (1987) Neurophysiology of rheumatic pain. *Baillière's Clinical Rheumatology*. 1 (1), pp. 1-26.

Hawker, G.A., Mian, S., Kendzerska, T. and French, M. (2011) Measures of adult pain visual analog scale for pain (VAS pain), numeric rating scale for pain (NRS pain), McGill pain questionnaire (MPQ), short-form McGill pain questionnaire (SF-MPQ), chronic pain grade scale (CPGS), short form-36 bodily pain scale (SF-36 BPS), and measure of intermittent and constant osteoarthritis pain (ICOAP). *Arthritis Care and Research*. 63 (Suppl 11), pp. s240-s252.

Hazes, J.M., Hayton, R., Burt, J. and Silman, A.J. (1994) Consistency of morning stiffness: an analysis of diary data. *British Journal of Rheumatology*. 33 (6), 562-565.

Hazes, J.M., Hayton, R. and Silman, A.J. (1993) A re-evaluation of the symptom morning stiffness. *The Journal of Rheumatology*. 20 (7), pp. 138-142.

Heiberg, T. and Kvien, T.K. (2002) Preferences for improved health examined in 1,024 patients with rheumatoid arthritis: pain has highest priority. *Arthritis Care and Research*. 47 (4), pp. 391-397.

Helliwell, P. (1995) The semiology of arthritis: discrimination between patients of the basis of their symptoms. *Annals of the Rheumatic Diseases*. 54, pp. 924-926.

Helliwell, P. (2004) Central nervous system involvement in perceived joint stiffness. *Rheumatology*. 43, pp. 255.

Helliwell, P., Howe, A. and Wright, V. (1987a) Functional assessment of the hand: reproducibility, acceptability, and utility of a new system for measuring strength. Annals of the Rheumatic Diseases. 46, pp. 203-208.

Helliwell, P., Howe, A. and Wright, V. (1987b) The measurement of stiffness in the rheumatoid hand. *Journal of Engineering in Medicine*. 16, pp. 203-207.

Helliwell, P., Howe, A. and Wright, V. (1988) Lack of objective evidence of stiffness in rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 47, pp. 754-758.

Helliwell, P., Smeathers, J.E. and Wright. (1995) The contribution of different tissues to stiffness of the wrist joint. Proceedings of the Institution of Mechanical Engineers, Part H. *Journal of Engineering in Medicine*. 208 (4), pp. 223-228.

Helliwell, P., Woodburn, J., Redmond, A., Turner, D. and Davys, H. (2007) *The Foot and Ankle in Rheumatoid Arthritis. A comprehensive guide.* London: Elsevier.

Helliwell, T., Brouwer, E., Pease, C.T., Hughes, R., Hill, C.L., Neill, L.M., Halls, S., Simon, L.S., Mallen, C.D., Boers, M., Kirwan, J.R. and Mackie, S.L. (2016) Development of a provisional core domain set for polymyalgia rheumatica: report from the OMERACT 12 Polymyalgia Rheumatica Working Group. *The Journal of Rheumatology*. 43 (1), pp. 182-186.

Hewlett, S., Ambler, N., Almeida, C., Blair, P.S., Choy, E., Dures, E., Hammond, A., Hollingworth, W., Kirwan, J., Plummer, Z., Rooke, C., Thorn, J., Tomkins, K. and Pollock, J. on behalf of the RAFT study team. (2015) Protocol for a randomised controlled trial for Reducing Arthritis Fatigue by clinical Teams (RAFT) using cognitive–behavioural approaches. *The British Medical Journal Open* [online]. Available from http://bmjopen.bmj.com/content/5/8/e009061.full [Accessed 22 January 2016].

Hewlett, S., Ambler, N., Almeida, C., Cliss, A., Hammond, A., Kitchen, K., Knops, B., Pope, D., Spears, M., Swinkels, A. and Pollock, J. (2011b) Self-management of fatigue in rheumatoid arthritis: a randomised controlled trial of group cognitive-behavioural therapy. *Annals of the Rheumatic Diseases*. 70, pp. 1060-1067.

Hewlett, S., Carr, M., Byron, M., Ryan, S., Kirwan, J., Richards, P., Carr, A. and Hughes, R. (2005a) Outcomes generated by patients with rheumatoid arthritis: How important are they? *Musculoskeletal Care*. 3, pp. 131-142.

Hewlett, S., Chalder, T., Choy, E., Cramp, F., Davis, B., Dures, E., Nicholls, C. and Kirwan, J. (2011a) Fatigue in rheumatoid arthritis: time for a conceptual model. *Rheumatology*. 50, pp. 1004-1006.

Hewlett, S., Cockshott, Z., Byron, M., Kitchen, K., Tipler, S., Pope, D. and Hehir, M. (2005b) Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. *Arthritis Care and Research*. 53 (5), pp. 697-702.

Hewlett, S., de Wit, M., Richards, P., Quest, E., Hughes, R., Heiberg, T. and Kirwan, J. (2006) Patients and professionals as research partners: challenges, practicalities, and benefits. *Arthritis Care and Research*. 55 (4), pp. 676-680.

Hewlett, S., Heir, M. and Kirwan, J.R. (2007) Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. *Arthritis Care and Research*. 57 (3), pp. 429-439.

Hewlett, S., Kirwan, J., Pollock, J., Mitchell, K., Hehir, M., Blair, P.S., Memel, D. and Perry, M.G. (2005c) Patient initiated outpatient follow up in rheumatoid arthritis: six year randomised controlled trial. *The British Medical Journal Open* [online]. Available from http://www.bmj.com/content/330/7484/171.full.pdf+html [Accessed 31 July 2012].

Hewlett, S., Nicklin, J., Bode, C., Carmona, L., Dures, E., Engelbrecht, M., Hagel, S., Kirwan, J,, Molto, A., Redondo, M. and Gossec, L. (2016a) Translating patient reported outcome measures: methodological issues explored using cognitive interviewing with three rheumatoid arthritis measures in 6 European languages. *Rheumatology*. 55, pp. 1009-1016.

Hewlett, S., Nicklin, J., Bode, C., Cramp, F., Carmona, L., Davis, B., Dures, E., Engelbrecht, M., Fransen, J., Gossec, L., Greenwood, R., Hagel, S., van de Laar, M., Petersson, I., Redondo, M., Molto, A., Schett, G. and Kirwan, J. (2016b) The revised Bristol RA Fatigue NRS, containing new NRS-coping anchors (BRAF-NRS V2 revised): validity across 6 European countries. *Annals of the Rheumatic Diseases*. 75 (Suppl 2), pp. 423.

Hewlett, S., Sanderson, T., May, J., Alten, R., Bingham, C.O., Cross, M., March, L., Pohl, C., Woodworth, T. and Bartlett, S.J. (2012) 'I'm hurting, I want to kill myself': rheumatoid arthritis flare is more than a high joint count-an international patient perspective on flare where medical help is sought. *Rheumatology*. 51 (1), pp. 69-76.

Hill, J. (2006) Pain and Stiffness. In: Hill, A. ed. (2006) *Rheumatology nursing. A creative approach.* Chichester: John Wiley & Sons Ltd, pp. 217-243.

Hjermstad, M.J., Fayers, P.M., Haugen, D.F., Caraceni, A., Hanks, G.W., Loge, J.H., Fainsinger, R., Aass, N. and Kaasa, S. on behalf of the European Palliative Care Research Collaborative (EPCRC) (2011). Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *Journal of Pain and Symptom management*. 41 (6), pp. 1073-1093.

Holloway, I. and Todres, L. (2003) The status of method: flexibility, consistency and coherence. *Qualitative Research*. 3 (3), pp. 345-357.

Houssien, D.A. McKenna, S.P. and Scott, D.L. (1997) The Nottingham Health Profile as a measure of disease activity and outcome in rheumatoid arthritis. *British Journal of Rheumatology*. 36, pp. 69-73.

Howe, A., Thompson, D. and Wright, V. (1985) Reference values for metacarpophalangeal joint stiffness in normal. *Annals of the Rheumatic Diseases*. 44, pp. 469-476.

Hughes, R.A., Mackie, S., Kirwan, J.R., Pease, C.T., Walsh, M. and Morris, M. (2012) What do patients with polymyalgia rheumatica mean when they describe stiffness? A qualitative study. *Arthritis and Rheumatism* [online] Available from: http://www.blackwellpublishing.com/acrmeeting/abstract.asp?MeetingID=789&id=102654 [Accessed 12 December 2012].

Hunt, S.M. and McEwen, J. (1980) The development of a subjective health indicator. *Social Health and Illness*. 2, pp. 231-246.

Hurkmans, E.J., van der Giesen, F.J., Bloo, H., Boonman, D.C., van der Esch, M., Fluit, M., Hilberdink, W.K., Peter, W.F., van der Stegen, H.P., Veerman, E.A., Verhoef, J., Vermeulen, H.M., Hendriks, H.M., Schoones, J.W., Vliet Vlieland, T.P. and Guideline Steering Committee. (2011) Physiotherapy in rheumatoid arthritis: development of a practice guideline. *Acta Rheumatologica Portuguesa*. 36 (2), pp. 146-158.

Hurkmans, E.J., Li, L., Verhoef, J. and T. P. M. Vliet Vlieland, T.P.M. (2012) Physical therapists' management of rheumatoid arthritis: results of a Dutch survey. *Musculoskeletal Care*. 10, pp. 142-148.

Huskinson, E.C. (1974) Measurement of pain. *The Lancet*. 2, pp. 1127-1131.

Hutchenson, G.D. and Sofroniou, N. (1999) *The Multivariate Social Scientist Introductory Statistics Using Generalized Linear Models*. London: SAGE Publications Ltd.

Hyrich, K, Symmons, D., Watson, K. and Silman, A. (2006) Baseline comorbidity levels in biologic and standard DMARD treated patients with rheumatoid arthritis: results from a national patient register. *Annals of the Rheumatic Diseases*. 65, pp. 895-898.

Ingpen, M.L. (1968) The quantitative measurement of joint changes in rheumatoid arthritis. *Annals of Physical Medicine*. 9 (8), pp. 322-327.

Ingpen, M.L. and Kendall, P.H. (1968) A simple apparatus for assessment of stiffness. *Annals of Physical Medicine*. 9 (5), pp. 203-205.

INVOLVE. (2012) *INVOLVE briefing notes for researchers: public involvement in NHS, public health and social care research*. Available from: http://www.invo.org.uk/wp-content/uploads/2014/11/9938_INVOLVE_Briefing_Notes_WEB.pdf [Accessed 21 November 2012].

Isik, A., Koca, S.S., Ozturk, A. and Mermi, O. (2007) Anxiety and depression in patients with rheumatoid arthritis. *Clinical Rheumatology*. 26, pp. 872-878.

Jaatun, E.A.A., Haugen, D.F., Dahl, Y. and Kofod-Petersen, A. (2015) Designing a reliable pain drawing tool: avoiding interaction flaws by better tailoring to patients' impairments. *Personal and Ubiquitous Computing*. 19, pp. 635-648.

Jastrząbek, R., Straburzyńska-Lupa, A., Rutkowski, R. and Romanowski W. (2013) Effects of different local cryotherapies on systemic levels of TNF-a, IL-6, and clinical parameters in active rheumatoid arthritis. *Rheumatology International*. 33 (8), pp. 2053-2060.

Johnson, R.B. and Onwuegbuzie, A.J. (2004) Mixed methods research: a research paradigm whose time has come. *Educational Researcher*. 33 (7), 14-26.

Johnson, R.B., Onwuegbuzie, A.J. and Turner, L.A. (2007) Toward a definition of mixed methods research. *Journal of Mixed Methods Research*. 1 (2), pp. 112-133.

Jones, P.W., Ziade, M.F.M., Davis, M.J. and Dawes, P.T. (1993) An index of disease activity in rheumatoid arthritis. *Statistics in Medicine*. 12, pp. 1171-1181.

Kalyoncu, U., Dougados, M., Daures, J-P. and Gossec, L. (2009) Reporting of patient-reported outcomes in recent trials in rheumatoid arthritis: a systematic literature review. *Annals of the Rheumatic Diseases*. 68, pp. 183-190.

Kay, J. and Upchurch, K.S. (2012) ACR/EULAR 2010 rheumatoid arthritis classification criteria. *Rheumatology*. 51 (Suppl 6), pp. vi5-vi9.

Kerr, C., Nixon, A. and Wild, D. (2010) Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcome research. *Expert Reviews of Pharmacoeconomics and Outcomes Research*. 10 (3), pp. 269-281.

Kett, C., Flint, J., Openshaw, M., Raza, K. and Kumar, K. (2010) Self-Management strategies used during flares of rheumatoid arthritis in an ethnically diverse population. *Musculoskeletal Care*. 8, pp. 204-214.

Khan, N.A., Yazici, Y., Calvo-alen, J., Dadoniene, J., Gossec, L., Hansen, T.M., Huisman, M., Kallikorm, R., Muller, R., Liveborn, M., Oding, R., Luchikhina, E., Naranjo, A., Rexhepi, S., Taylor, P., Tlustochowich, W., Tsirogianni, A. and Sokka, T. (2009) Reevaluation of the role of duration of morning stiffness in the assessment of rheumatoid arthritis activity. *The Journal of Rheumatology*. 36 (11), pp. 2435-2442.

Kim, J. and Mueller, C.W. (1978) *Factor Analysis: Statistical Methods and Practical Issues*. London: SAGE Publications Ltd.

Kincaid, J.P., Fishburne, R.P., Rogers, R.L. and Chissom, B.S. (1975) Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. *Research Branch Report 8-75. Chief of Naval Technical Training: Naval Air Station Memphis* [online]. Available from: http://www.dtic.mil/dtic/tr/fulltext/u2/a006655.pdf [Accessed 14 October 2015].

Kirwan, J.R. (2011) Targeting the time of day for glucocorticoid delivery in rheumatoid arthritis. *International Journal of Clinical Rheumatology*. 6 (3), pp. 273-279.

Kirwan, J.R., Chaput de Saintonge, D.M., Joyce, C.R.B. and Currey, H.L.F. (1984) Clinical judgements in rheumatoid arthritis III. British rheumatologists' judgements of 'change in response to therapy'. *Annals of the Rheumatic Diseases*. 43, pp. 686-694.

Kirwan, J.R., Heiberg, T., Hewlett, S., Hughes, R., Kvien, T., Ahlmen, M., Boers, M., Minnock, P., Saag, K., Shea, B., Suarez Almazor, M. and Taal, E. (2003) Outcomes from the patient perspective workshop at OMERACT 6. *The Journal of Rheumatology.* 30 (4), 868-872.

Kirwan, J.R., Minnock, P., Adebajo, A. Bresnihan, B., Choy, E, de Wit, M., Hazes, M., Richards, P., Sagg, K., Suarez-Almazor, M, Wells, G. and Hewlett, S. (2007) Patient perspective: fatigue as a recommended patient centred outcome measure in rheumatoid arthritis. *The Journal of Rheumatology*. 34 (5), pp. 1174-1177.

Kitzinger, J. (1994) The methodology of focus groups: the importance of interaction between research participants. *Sociology of Health and Illness*. 16 (1), pp. 103-121.

Kline, P. (2000) *The Handbook of Psychological Testing*. London: Routledge.

Koch, T. and Harrington, A. (1998) Reconceptualising rigour: the case for reflexivity. *Journal of Advanced Nursing*. 28 (4), pp. 882-890.

Krishnan, E., Lingala, B., Bruce, B. and Fries, J.F. (2011) Disability in rheumatoid arthritis in the era of biological treatments. *Clinical and Epidemiological Research*. 71, pp. 213-218.

Krueger, R.A. and Casey, M.A. (2009) *Focus Groups: A Practical Guide for Applied Research*. London: SAGE Publications Ltd.

Kvien, T.K., Glennås, A., Knudsrød, O.G. and Smedstad, L.M. (1996) The validity of self-reported diagnosis of rheumatoid arthritis: results from a population survey followed by clinical examinations. *Journal of Rheumatology.* 23 (11), pp. 1866-1871.

Labitigan, M., Sharaf, P.I.I., Swearingen, C. and Yazici, Y. (2010) Rheumatoid arthritis patients who participated in randomized clinical trials in 1980s versus 2000s: clinical implications of similarities and differences. *Annals of the Rheumatic Diseases* [online]. Available from: http://www.abstracts2view.com/eular/lookup_view.php?page=3&start=40&stop=20&word=Swearingen&where=authors&return=%2Feular%2Fauthorindex.php%3Fnum%3D18%26stop%3D39%26page%3D91%26start%3D3510 [Accessed 22 June 2015].

Lambert, S.D. and Loiselle, C.G. (2007) Combining individual interviews and focus groups to enhance data richness. *Journal of Advanced Nursing*. 62 (2), pp. 228-237.

Lansbury, J. (1956) Quantification of the activity of rheumatoid arthritis. 2. Recession of morning stiffness as patients go into remission. *The American Journal of the Medical Sciences*. 232 (1), pp. 8-11.

Lansbury, J. (1958) Report of a three-year study on the systemic and articular indexes in rheumatoid arthritis. *Arthritis and Rheumatism*. 1 (6), pp. 505-522.

Larkin, M., Watts, S. and Clifton, E. (2006) Giving voice and making sense in interpretative phenomenological analysis. *Qualitative Research in Psychology*. 3 (2), pp. 102-120.

Leeb, B.F. and Bird, H.A. (2004) A disease activity score for polymyalgia rheumatica. *Annals of the Rheumatic Diseases.* 63, pp. 1279-1283.

Leeb, B.F., Haindl, P.M., Maktari, A., Nothnagl, T. and Rintelen, B. (2008) Patient-centered rheumatoid arthritis disease activity assessment by a modified RADAI. *The Journal of Rheumatology.* 35 (7), pp. 1294-1299.

Leeb, B.F., Sautner, J., Andel, I. and Rintelen, B. (2003) SACRAH: a score for assessment and quantification of chronic rheumatic affections of the hands. *Rheumatology.* 42, pp. 1173-1178.

Leech, N.L. and Onwuegbuzie, A.J. (2009) A typology of mixed methods research designs. *Quality and Quantity*. 43 (2), pp. 265-275.

Legard, R., Keegan, J. and Ward, K. (2003) In-depth interviews. *Qualitative Research Practice. A Guide for Social Science Students and Researchers*. London: SAGE Publications Ltd, pp. 138-169.

Leidy, N.K. and Vernon, M. (2008) Perspectives on patient-reported outcomes content validity and qualitative research in a changing clinical trial environment. *Pharmacoeconomics*. 26 (5), pp. 363-370.

Leung, A.M.H., Farewell, D., Lau, C.S. and Choy, E. (2012) Defining criteria for rheumatoid arthritis patient-derived disease activity score that corresponds to disease activity score 28 and clinical disease activity index based statuses and response criteria. *Arthritis and Rheumatism*. 64 (Suppl 10), pp. 393.

Levin, R.W. Park, J., Ostrov, B., Reginato, A., Baker, D.G., Bomalaski, J.S., Borofsky, M., Gardiner, M., Leventhal, L., Louthrenoo, W., von Feldt, J., Kolasinski, S. and Schumacher, H.R. (1996) Clinical assessment of the 1987 American College of Rheumatology criteria for rheumatoid arthritis. *Scandinavian Journal of Rheumatology*. 25 (5), pp. 277-281.

Lewis, J. (2003) Design issues. In Ritchie, J. And Lewis, J. eds. (2003) *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: SAGE Publications Ltd, pp. 47-76.

Lie, E., Woodworth, T.G., Christensen, R., Kvien, T., Bykerk, V., Furst, D.E., Bingham, C.O., Choy, E.H. and the OMERACT RA Flare Working Group. (2014) Validation of OMERACT preliminary rheumatoid arthritis flare domains in the NOR-DMARD study. *Annals of the Rheumatic Diseases*. 73, pp. 1781-1787.

Lineker, S.A., Badley, E., Charles, C., Hart, L. and Streiner, D. (1999) Defining morning stiffness in rheumatoid arthritis. *The Journal of Rheumatology*. 26 (5), pp. 1052-1057.

Linting, M. (2015) Email to Serena Halls, 29 June.

Linting, M. Meulman, J.J. Groenen, P.J.F. and van der Kooij, A.J. (2007a) Nonlinear principal components analysis: introduction and application. *Psychological Methods*. 12 (3), pp. 336-358.

Linting, M. Meulman, J.J. Groenen, P.J.F. and van der Kooij, A.J. (2007b) Stability of nonlinear principal components analysis: an empirical study using the balanced bootstrap. *Psychological Methods*. 12 (3), pp. 359-379.

Linting, M. and van der Kooij, A. (2012) Nonlinear principal component analysis with CATPCA: a tutorial. *Journal of Personality Assessment*. 94 (1), pp. 12-25.

Lopez-Olivo, M.A., Amezaga Urruela, M., McGahan, L., Pollono, E.N. and Suarez-Almazor, M.E. (2015) Rituximab for rheumatoid arthritis. *Cochrane Database of Systematic Reviews* [online]. Available from: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007356.pub2/full [Accessed 23 November 2015].

Lorenzo-Seva, U. and Ferrando, P.J. (2013) Manual of the program FACTOR v.9.20. [computer programme].

Lorenzo-Seva, U. and Ferrando, P.J. (2014) POLYMAT-C: a comprehensive SPSS program for computing the polychoric correlation matrix. *Behavioural Research Methods*. 47 (3), pp. 884-889.

Lukas, C., Landewe, R., Sieper, J., Dougados, M., Davis, J., Braun, J., van der Linden, S. and van der Heijde, S. (2009) Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Annals of the Rheumatic Diseases*. 68 (1), pp. 18-24.

Luqmani, R., Hennell, S., Estrach, C., Basher, D., Birrell, F., Bosworth, A., Burke, F., Callaghan, C., Candal-Couto, J., Fokke, C., Goodson, N., Homer, D., Jackman, J., Jeffreson, P., Oliver, S., Reed, M., Sanz, L., Stableford, Z., Taylor, P., Todd, N., Warburton, L., Washbrook, C. and Wilkinson, M. on behalf of the British Society for Rheumatology and British Health Professionals in Rheumatology Standards, Guidelines and Audit Working Group. (2009) British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of rheumatoid arthritis (after the first 2 years). *Rheumatology*. 48 (4), pp. 436-439.

Luqmani, R., Hennell, S., Estrach, C., Birrell, F., Bosworth, A., Davenport, G., Fokke, C., Goodson, N., Jeffreson, P., Lamb, E., Mohammed, R., Oliver, S., Stableford, Z., Walsh, D., Washbrook, C. and Webb, F. on behalf of the British Society for Rheumatology and British Health Professionals in Rheumatology Standards, Guidelines and Audit Working Group. (2006) British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of rheumatoid arthritis (the first 2 years). *Rheumatology*. 45 (9), pp. 1167-1169.

MacCallum, R.C., Widaman, K.F., Zhang, S. and Hong, S. (1999) Sample size in factor analysis. *Psychological Methods*. 4 (1), pp. 84-99.

MacKay, C., Badley, E.M., Jaglal, S.B., Sale, J. and Davis, A.M. (2014b) “We’re all looking for solutions”: A qualitative study of the management of knee symptoms. *Arthritis Care and Research*. 66 (7), pp. 1033-1040.

MacKay, C., Jaglal, S.B., Sale, J., Badley, E.M. and Davis, A.M. (2014a) A qualitative study of the consequences of knee symptoms: ‘It’s like you’re an athlete and you go to a couch potato’. *The British Medical Journal Open* [online]. Available from: http://bmjopen.bmj.com/content/4/10/e006006.full.pdf+html [Accessed 4 February 2015].

Mackie, S.L., Arat, S., da Silva, J., Duarte, C., Halliday, S., Hughes, R., Morris, M., Pease, C.T., Sherman, J.W., Simon, L.S., Walsh, M., Westhovens, R., Zakout, S. and Kirwan, J. (2014) Polymyalgia Rheumatica (PMR) Special Interest Group at OMERACT 11: outcomes of importance for patients with PMR. *The Journal of Rheumatology*. 41 (4), pp. 819-823.

Mackie, S.L., Hughes, R., Walsh, M., Day, J., Newton, M., Pease, C., Kirwan, J. and Morris, M. (2015) “An impediment to living life”: why and how should we measure stiffness in polymyalgia rheumatica? *PLoS ONE* [online]. Available from: http://www.plosone.org/article/fetchObject.action?uri=info:doi/10.1371/journal.pone.0126758&representation=PDF [Accessed 26 May 2015].

Mallya, R.K. and Mace, B.E.W. (1981) The assessment of disease activity in rheumatoid arthritis using a multivariate analysis. *Rheumatology and Rehabilitation*. 20, pp. 14-17.

Margollis, R.B., Chibnall, J.T. and Tait, R.C. (1988) Test-retest of the pain drawing instrument. *Pain*. 33 (1), pp. 59-51.

Margollis, R.B., Tait, R.C. and Krause, S.J. (1986) A rating system for use with patient pain drawings. *Pain*. 24 (1), pp. 57-65.

Maska, L., Anderson, J. and Michaud, K. (2011) Measures of functional status and quality of life in rheumatoid arthritis. *Arthritis Care and Research*. 63 (Suppl 11), pp. s4-s13.

Mason, J.H., Anderson, J.J., Meenan, R.F., Haralson, K.M., Lewis-Stevens, D. and Kaine, J.L. (1992) The rapid assessment of disease activity in rheumatology (RADAR) questionnaire. *Arthritis and Rheumatism*. 35, pp. 156-162.

Matteson, E.L., Maradit-Kremers, H., Cimmino, M.A., Schmidt, W.A., Schrmer, M., Salvarani, C., Bachta, A., Dejaco, C., Duftner, C., Slott-Jensen, H., Poor, G., Kaposi, N.P., Mandl, P., Balint, P.V., Schmidt, Z., Iagnocco, A., Cantini, F., Nannini, C., Macchioni, P., Pipitone, N., del Amo, M., Espigol-Frigole, G., Cid., M.C., Martinez-Taboada, V.M., Nordborg, E., Direskeneli, H., Aydin, S.Z., Ahmed, K., Hazelman, B., Pease, C., Wakefield, R.J., Luqmani, R., Abril, A., Marcus, R., Gonter, N.J., Maz, M., Crowson, C.S. and Dasgupta, B. (2012) Patient-reported outcomes in polymyalgia rheumatica. *The Journal of Rheumatology*. 39 (4), pp. 795-803.

Mattila, K., Buttgereit, F. and Tuominen, R. (2014) Impact of morning stiffness on working behaviour and performance in people with rheumatoid arthritis. *Rheumatology International*. 34 (12), pp. 1751-1758.

Mays, N. and Pope, C. (1995) Rigour and qualitative research. *British Medical Journal*. 311, pp. 109-112.

Mays, N. and Pope, C. (2000) Qualitative research in health care: Assessing quality in qualitative research. *British Medical Journal*. 320, 50-52.

McConnell, S., Kolopack, P. and Davis, A.M. (2001) The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): A review of its utility and measurement properties. *Arthritis Care and Research*. 45, pp. 453-461.

McCormack, H.M., Horne, D.J., Sheather, S. (1988) Clinical applications of visual analogue scales: a critical review. *Psychological Medicine*. 18 (4), pp. 1007-1019.

McInnes, I.B. and Schett, G. (2007) Cytokines in the patient with rheumatoid arthritis. *Nature Reviews: Immunology*. 7, pp. 429-442.

McLaughlin, G.H. (1969) SMOG grading – a new readability formula. *Journal of Reading*. 12 (8), pp. 639-646.

McLeod, J. (2001) *Qualitative Research in Counselling and Psychotherapy*. London: SAGE Publications Ltd.

Meenan, R.F., Gertman, P.M. and Mason, J.H. (1980) Measuring health status in arthritis: the arthritis impact measurement scales. *Arthritis and Rheumatism*. 23, pp. 146-152.

Meenan, R.F., Mason, J.H., Anderson, J.J., Guccione, A.A. and Kazis, L.E. (1992) AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. *Arthritis and Rheumatism*. 35 (1), pp. 1-10.

Melzack, R. and Wall, P.D. (1965) Pain mechanisms: a new theory. *Science*. 150 (3699), pp. 971-979.

Mengi, G., Gogus, F. and Beyazova, M. (2014) Is it possible to quantify morning stiffness objectively? *Annals of the Rheumatic Diseases* [online]. Available from: http://www.abstracts2view.com/eular/view.php?nu=EULAR14L_SAT0147 [Accessed 17 December 2014].

Mengshoel, A.M. (2008) Living with a fluctuating illness of ankylosing spondylitis: a qualitative study. *Arthritis Care and Research*. 59 (10), pp. 1439-1444.

Mertens, M. and Singh, J.A. (2009) Anakinra for rheumatoid arthritis. *Cochrane Database of Systematic Reviews* [online]. Available from: http://www.cochrane.org/CD005121/MUSKEL_anakinra-for-rheumatoid-arthritis [Accessed 23 November 2015].

Meyer, T. (2007) Problems completing questionnaires on health status in medical rehabilitation patients. *Journal of Rehabilitation Medicine*. 39, pp. 633-639.

Meyrick, J. (2006) What is good qualitative research? A first step towards a comprehensive approach to judging rigour/quality. *Journal of Health Psychology*. 11 (5), pp. 799-808.

Miller, D. and Robert, L. (1988) Invariance of item characteristic functions with variations in instructional coverage. *Journal of Educational Measurement.* 25 (3), pp. 205-219.

Miller, G.A. (1994) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*. 101 (2), pp. 343-352.

Minnock, P., FitzGerald, O. and Bresnihan, B. (2003) Women with established rheumatoid arthritis perceive pain as the predominant impairment of health status. *Rheumatology*. 42, pp. 995-1000.

Minnock, P, Bresnihan, B. (2004). Pain outcome and fatigue levels reported by women with established rheumatoid arthritis. *Arthritis and Rheumatism*. 50 (Suppl 9), pp. 1198.

Minnock, P. and Bresnihan, B. (2008) Fatigue is related to poor pain outcomes in women with established rheumatoid arthritis. *Clinical Experimental Rheumatology*. 26 (4), 707-708.

Mok, C.C., Cha, H.S., Hidayat, R., Nguyen, L.T.N., Perez, E.C., Ramachandran, R., Tsay, G.J., Yoo, D.H. and for the Asia Pacific Morning Stiffness in Rheumatoid Arthritis Expert Panel. (2015) The importance of assessment and management of morning stiffness in Asian patients with rheumatoid arthritis: recommendations from an expert panel. *International Journal of Rheumatic Diseases*. 19 (1), pp. 30-37.

Morgan, D.L. (1997) *Focus Groups as Qualitative Research*. London: SAGE Publications Ltd.

Morris, R. and Carstairs, V. (1991) Which deprivation? A comparison of selected deprivation indexes. *Journal of Public Health Medicine*. 13 (4), pp. 318-326.

Murtagh, F.E.M., Addington-Hall, J.M. and Higginson, I.J. (2007) The value of cognitive interviewing techniques in palliative care research. *Palliative Medicine*. 21, pp. 87-93.

Nasir, S.H. Troynikov, O. and Massy-Westropp, N. (2014) Therapy gloves for patients with rheumatoid arthritis: a review. *Therapeutic Advances in Musculoskeletal Disease.* 6 (6), pp. 226-237.

Newington, L. and Metcalfe, A. (2014) Factors influencing recruitment to research: qualitative study of the experiences and perceptions of research teams. *BMC Medical Research Methodology* [online]. Available from: http://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-14-10 [Accessed 17 April 2015].

National Institute for Health and Care Excellence. (2009) *Rheumatoid Arthritis in Adults: Management*. Report number: CG79. London: National Institute for Health and Care Excellence.

Navarro-Sarabia, F., Ariza-Ariza, R., Hernandez-Cruz, B. and Villanueva, I. (2005) *Cochrane Database of Systematic Reviews* [online]. Available from: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD005113.pub2/full [Accessed 23 November 2015].

Naz, S.M. and Symmons, D.P.M. (2007) Mortality in established rheumatoid arthritis. *Best Practice and Research Clinical Rheumatology*. 21 (5), pp. 871-883.

Ndosi, M., Lewis, M., Hale, C., Quinn, H., Ryan, S., Emery, P, Bird, H. and Hill, J. (2014) The outcome and cost-effectiveness of nurse-led care in people with rheumatoid arthritis: a multicentre randomised controlled trial. *Annals of the Rheumatic Diseases*. 73, pp. 1975-1982.

Nelson, E.C., Eftimovska, E., Lind, C., Hager, A., Wasson, J.H. and Lindblad, S. (2015) Patient reported outcome measures in practice. *The British Medical Journal* [online]. Available from: http://www.bmj.com/content/350/bmj.g7818 [Accessed 7 October 2015].

Nicklin, J. (2009) *The development of scales to measure fatigue in people with rheumatoid arthritis*. PhD, University of the West of England.

Nicklin, J., Cramp, F., Kirwan, J., Greenwood, R., Urban, M. and Hewlett, S. (2010b) Measuring fatigue in rheumatoid arthritis: A cross-sectional study to evaluate the Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional Questionnaire, Visual Analogue & Numerical Rating Scales. *Arthritis Care and Research*. 62 (11), pp. 1559-1568.

Nicklin, J., Cramp, F., Kirwan, J., Urban, M. and Hewlett, S. (2010a) Collaboration with patients in the design of patient-reported outcome measures: Capturing the experience of fatigue in rheumatoid arthritis. *Arthritis Care and Research*. 62 (11), pp. 1552-1558.

Nicklin, J., Hewlett, S.E., Two, R., Kirwan, J.R. (2014) Have you felt embarrassed or upset? An example of the care needed in translating patient reported outcome measures. *Annals of the Rheumatic Diseases*. 73 (Suppl 2), pp. 94.

Nikolaus, S., Bode, C., Taal, E. and van de Laar, M.A.F.J. (2013) Fatigue and factors related to fatigue in rheumatoid arthritis: a systematic review. *Arthritis Care and Research*. 65 (7), pp. 1128-1146.

Norquist, J.M., Girman, C., Fehnel, S., DeMuro-Mercon, C. and Santanello, N. (2012) Choice of recall period for patient-reported outcome (PRO) measures: criteria for consideration. *Quality of Life Research*. 21, pp. 1013-1020.

Nunnally, J.C. and Bernstein, I.H. (1994) *Psychometric Theory*. New York: McGraw-Hill.

Office of National Statistics. (2014) Internet access – households and individuals 2014. *Office of National Statistics* [online]. Available from: http://www.ons.gov.uk/ons/dcp171778_373584.pdf [Accessed 1 October 2015].

Oliver, S. (2006) The immune system and rheumatic disease. In: Hill, J. ed (2006) *Rheumatology Nursing: A Creative Approach*. Chichester: John Wiley & Sons Ltd, pp. 95-122.

Olsson, U. (1979). Maximum likelihood of the polychoric correlation coefficient. *Psychometrika*. 44, pp. 443-460.

O'Neill, J., Rader, T., Guillemin, F., Boonen, A., Christensen, R., Lyddiatt, A., Pardo Pardo, J., Welch, V., Singh, J.A. and Tugwell, P. (2014) Including health equity considerations in development of instruments for rheumatology research: an introduction to a novel OMERACT paradigm. *The Journal of Rheumatology*. 41 (1), pp. 150-152.

Onwuegbuzie, A.J. and Leech, N.L. (2007) A call for qualitative power analyses. *Quality and Quantity*. 41, pp. 105-121.

Oppenheim, A.N. (1992) *Questionnaire Design, Interviewing and Attitude Measurement*. London: Continuum, Pinter Publications.

Orbai, A-M., Halls, S., Hewlett, S., Bartlett, S.J., Leong, A.L. and Bingham, C.O. (2015) More than just minutes of stiffness in the morning: report of the rheumatoid arthritis flare group on stiffness breakout sessions. *The Journal of Rheumatology*. 42 (11), pp. 2182-2184.

Orbai, A-M., Smith, K.C., Bartlett, S.J., de Leon, E. and Bingham, C.O. (2014) "Stiffness has different meanings, I think, to everyone". Examining stiffness from the perspective of people living with rheumatoid arthritis. *Arthritis Care and Research.* 66 (11), pp. 1662-1672.

O'Reilly, M. and Parker, N. (2013) 'Unsatisfactory saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research*. 13 (2), pp. 190-197.

Ortony, A. (1975) Why metaphors are necessary and not just nice. *Educational Theory*. 25 (1), pp. 45-53.

Overman, C.L., Jurgens, M.S., Bossema, E.R., Jacobs, J.W.G., Bijlsma, J.W.J. and Geenen, R. (2014) Change of psychological distress and physical disability with rheumatoid arthritis over the last two decades. *Arthritis Care and Research*. 66 (5), pp. 671-678.

Overman, C.L., Kool, M.B. da Silva, J.A.P. and Geenen, R. (2015) The prevalence of severe fatigue in rheumatic diseases: an international study. *Clinical Rheumatology* [online]. Available from: http://link.springer.com/article/10.1007%2Fs10067-015-3035-6 [Accessed 22 January 2016].

Pallant, J. (2010) *SPSS Survival Manual. A Step by Step Guide to Data Analysis Using the SPSS Program*. Maidenhead: McGraw Hill.

Patrick, D.L., Burke, L., Gwaltney, C.J., Leidy, N.K., Martin, M.L., Molsen, E. and Ring, L. (2011a) Content validity - establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 1 – eliciting concepts for a new PRO instrument. *Value in Health*. 14 (8), pp. 967-77.

Patrick, D.L., Burke, L., Gwaltney, C.J., Leidy, N.K., Martin, M.L., Molsen, E. and Ring, L. (2011b) Content validity - establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 2 – assessing respondent understanding. *Value in Health*. 14 (8), pp. 978-88.

Patrick, D.L., Burke, L., Powers, J.H., Scott, J.A., Rock, E.P., Dawisha, A., O'Neill, R. and Kennedy, D.L. (2007) Patient-reported outcomes to support medical product labeling claims: FDA perspective. *Value in Health*. 10 (Suppl 2), pp. s125-s137.

Patton, M.Q. (2002) *Qualitative Research and Evaluation Methods*. London: SAGE Publications Ltd.

Paulus, H.E., Egger, M.J., Ward, J.R., Williams, H.J. and the cooperative systematic studies of rheumatic disease group. (1990) Analysis of improvement in individual rheumatoid arthritis patients treated with disease modifying anti rheumatic drugs, based on the findings in patients treated with placebo. *Arthritis and Rheumatism*. 33, pp. 477-484.

Pedhazur, E and Schmelkin, L. (1991) *Measurement, Design and Analysis: An Integrated Approach*. Hove: Psychology Press, Taylor and Francis Group.

Perry, M.G., Kirwan, J.R., Jessop, D.S. and Hunt, L.P. (2009) Overnight variations in cortisol, interleukin 6, tumour necrosis factor alpha and other cytokines in people with rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 68 (1), pp. 63-68.

Pett, M.A. (1997) *Nonparametric Statistics for Healthcare Research*. London: SAGE Publications Ltd.

Pett, M.A., Lackey, N.R. and Sullivan, J.J. (2003) *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*. London: SAGE Publications Ltd.

Phillips, S. and Dow, L. (2012) Impact of impaired morning function on quality of life in rheumatoid arthritis: results of an exploratory patient survey. *International Journal of Clinical Rheumatology*. 7 (6), pp. 597-606.

Pinals, R.S., Masi, A.T., Larsen, R.A., and the subcommittee for criteria for remission in rheumatoid arthritis of the American Rheumatism Association diagnostic and therapeutic criteria committee. (1981) Preliminary criteria for clinical remission in rheumatoid arthritis. *Arthritis and Rheumatism*. 24 (10), pp. 1308-1315.

Pincus, T. (2005) The American College of Rheumatology (ACR) core data set and derivative “patient only” indices to assess rheumatoid arthritis. *Clinical and Experiential Rheumatology*. 23 (Suppl 39), pp. s109-s113.

Pincus, T., Callahan, L.F., Brooks, R.H., Fuchs, H.A., Olsen, N.J. and Kaye, J.J. (1989) Self-report questionnaire scores in rheumatoid arthritis compared with traditional physical, radiographic, and laboratory measures. *Annals of Internal medicine*. 110 (4), pp. 259-266.

Pincus, T., Furer, V., Keystone, E., Yazici, Y., Bergman, M.J. and Luijtens, K. (2008) RAPID3 (Routine Assessment of Patient Index Data 3) severity categories and response criteria: similar results to DAS28 (Disease Activity Score) and CDAI (Clinical Disease Activity Index) in the RAPID 1 (Rheumatoid Arthritis Prevention of Structural Damage) clinical trial of certolizumab pegol. *Arthritis Care and Research*. 63 (8), pp. 1142-1149.

Pincus, T., Summey, J.A., Soraci, S.A., Wallston, K.A. and Hummon, N.P. (1983) Assessment of patient satisfaction in activities of daily living using a modified Stanford health assessment questionnaire. *Arthritis and Rheumatism*. 26, pp. 1346-1353.

Polsky, D., Doshi, J.A., Marcus, S., Oslin, D., Rothbard, A., Thomas, N. and Thompson, C.L. (2005) Long-term risk for depressive symptoms after a medical diagnosis. *Archives of Internal Medicine*. 165, pp. 1260-1266.

Poole, J.L. (2011) Measures of hand function. *Arthritis Care and Research*. 63 (Suppl 11), pp. s189 -s199.

Prevoo, M.L.L., van’t Hof, M.A. Kuper, H.H., van Leeuwen, M.A., van de Putte, L.B.A. and van Riel, P.L.C.M. (1995) Modified disease activity scores that included twenty-eight joint counts: Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis and Rheumatism*. 38 (1), pp. 44-48.

QSR International Pty Ltd. (2012) Nvivo Qualitative Data Analysis Software (10) [computer programme].

Quadri, N., Langel, K., Muehlhausen, W., O’Donohoe, P. and Wild, D, (2012) Exploring patient perceptions of, and preferences for, pain response scales. *Value in Health*. 15 (7), pp. A482.

Rattray, J. and Jones, M.C. (2007) Essential elements of questionnaire design and development. *Issues in Clinical Nursing*. 16 (2), pp. 234-243.

Rhind, V.M. (1988) *The validity and inter-relationship of subjective and objective measurements of stiffness in human joints* [online]. MSc, Durham University. Available from: http://etheses.dur.ac.uk/6402/ [Accessed 23 February 2015].

Rhind, V.M., Unsworth, A. and Haslock, I. (1987) Assessment of stiffness in rheumatology: the use of rating scales. *Rheumatology*. 26 (2), pp. 126-130.

Richards, H. and Emslie, C. (2000) The ‘doctor’ or the ‘girl from the University’? Considering the influence of professional roles on qualitative interviewing. *Family Practice*. 17 (1), pp. 71-75.

Rintelen, B., Haindl, P.M., Mai, H.T.H., Sautner, J., Maktari, A. and Leeb, B.F. (2009) A tool for the assessment of hand involvement in rheumatic disorders in daily routine e the SF-SACRAH (short form score for the assessment and quantification of chronic rheumatic affections of the hands). *Osteoarthritis and Cartilage*. 17, pp. 59-63.

Ritchie, J. (2003) The Applications of qualitative methods to social research. In Ritchie, J. and Lewis, J. (2003) *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: Sage Publications Ltd, pp. 24-46.

Ryan, S., Lillie, K., Thwaites, C. and Adams, J. (2013) 'What I want clinicians to know' - experiences of people with arthritis. *British Journal of Nursing*. 22 (14), pp. 808-812.

Sage, N., Sowden, M., Chorlton, E. and Edeleanu, A. (2008) What is the cognitive behavioural approach? In: Sage, N., Sowden, M., Chorlton, E. and Edeleanu, A. eds. (2008) *Cognitive Behaviour Therapy for Chronic Illness and Palliative Care*. Chichester: John Wiley & Sons Ltd, pp. 3-9.

Salmon, V. (2015) *Development of a physical activity intervention for managing fatigue in rheumatoid arthritis*. PhD, University of the West of England.

Sandelowski, M. (1995) Focus on qualitative methods: sample size in qualitative research. *Research in Nursing and Health*. 18, pp. 179-183.

Sandelowski, M. (2000) Focus on research methods combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing and Health*. 23, pp. 246-255.

Sanderson, T.C. (2009) *Well-being and the prioritization of treatment outcomes in rheumatoid arthritis*. PhD, University of the West of England.

Sanderson, T.C., Hewlett, S.E., Flurey, C., Dures, E., Richards, P. and Kirwan J.R. (2011) The impact triad (severity importance, self-management as a method of enhancing measurement of personal life impact of rheumatic disease. *The Journal of Rheumatology*. 38 (2), pp. 191-194.

Sarzi-Puttini, P., Fiorini, T., Panni, B., Turiel, M., Cazzola, M. and Atzeni, F. (2002) Correlation of the score for subjective pain with physical disability, clinical and radiographic scores in recent onset rheumatoid arthritis. *BMC Musculoskeletal Disorders* [online]. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC117789/pdf/1471-2474-3-18.pdf [Accessed 24 June 2013].

Sautner, J., Andel, I., Rintelen, I. and Leeb, B.F. (2004) Development of the M-SACRAH, a modified, shortened version of SACRAH (score for the assessment and quantification of chronic rheumatoid affections of the hands). *Rheumatology*. 43, pp. 1409-1413.

Schmitt, T.A. (2011) Current methodological considerations in exploratory factor and confirmatory factor analysis. *Journal of Psychoeducational Assessment*. 29 (4), pp. 304-321.

Scientific Advisory Committee of the Medical Outcomes Trust. (2002) Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research*. 11, pp. 193-205.

Scott, D.L., Pugner, K., Kaarela, K., Doyle, D.V., Woolf, A., Holmes, J. and Hieke, K. (2000) The links between joint damage and disability in rheumatoid arthritis. *Rheumatology*. 39, pp. 122-132.

Scott, J.T. (1960) Morning stiffness in rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 19, pp. 361-368.

Seale, C. (1999) *The Quality of Qualitative Research*. London: Sage Publications Ltd.

Shirinsky, I., Polovnikova, O., Kalinovskaya, N. and Shirinsky, V. (2013) The effects of fenofibrate on inflammation and cardiovascular markers in patients with active rheumatoid arthritis: a pilot study. *Rheumatology International*. 33, pp. 3045-3048.

Siemons, L., ten Klooster, P.M., Taal, E., Kuper, I., van Riel, P.L.C.M., Glas, C.A.W. and van de Laar, M.A.F.J. (2013) Contribution of assessing forefoot joints in early rheumatoid arthritis patients: insights from item response theory. *Arthritis Care and Research*. 65 (2), pp. 212-219.

Sieper, J., Rudwaleit, M., Baraliakos, X., Brandt, J., Braun, J., Burgos-Vargas, R., Dougados, M., Hermann, K-G., Landewe, R., Maksymowych, W. and van der Heijde, D. (2009) The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Annals of the Rheumatic Diseases*. 68 (Suppl II), pp. ii1-ii44.

Silman, A.J. and Macfarlane, G.J. (2002) *Epidemiological Studies. A Practical Guide*. Cambridge: Cambridge University Press.

Silman, A.J. and Pearson, J.E. (2002) Epidemiology and genetics of rheumatoid arthritis. *Arthritis Research and Therapy*. 4 (Suppl 3), pp. s265-s272.

Smith, J.A., Flowers, P. and Larkin, M. (2009) *Interpretative Phenomenological Analysis: Theory, Method and Research*. London: SAGE Publications.

Smolen, J.S., Landewé, R., Breedveld, F.C., Dougados, M., Emery, P., Gaujoux-Viala, C., Gorter, S., Knevel, R., Nam, J., Schoels, M., Aletaha, D., Buch, M., Gossec, L., Huisinga, T., Bijlsma, J.W., Burmester, G., Combe, B., Cutolo, M., Gabay, C., Gomez-Reino, J., Kouloumas, M., Kvien, T.K., Martin-Mola, E., McInnes, I., Pavelka, K., van Riel, P., Scholte, M., Scott, D., Sokka, T., Valesini, G., van Vollenhoven, R., Winthrop, K.L., Wong, J., Zink, A. and van der Heijde, D. (2010) EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Annals of the Rheumatic Diseases*. 69, pp. 964-975.

Snow, S. and Kirwan, J.R. (1988) Visual analogue scales: a source of error. *Annals of the Rheumatic Diseases*. 47, pp. 526.

Sokka, T., Kautiainen, H., Pincus, T., Verstappen, S.M.M., Aggarwal, A., Alten, R., Andersone, D., Badsha, H., Baecklund, E., Belmonte, M., Craig-Müller, J., da Mota, L.M.H., Dimic, A., Fathi, N.A., Ferraccioli, G., Fukuda, W., Géher, P., Gogus, F., Hajjaj-Hassouni, N., Hamoud, H., Haugeberg, G., Henrohn, D., Horslev-Petersen, K., Ionescu, R., Karateew, D., Kuuse, R., Laurindo, I.M.M., Lazovskis, J., Luukkainen, R., Mofti, A., Murphy, E., Nakajima, A., Oyoo, O., Pandya, S.C., Pohl, C., Predeteanu, D., Rexhepi, M., Rexhepi, S., Sharma, B., Shono, E., Sibilia, J., Sierakowski, S., Skopouli, F.N., Stropuviene, S., Toloza, S., Valter, I., Woolf, A., Yamanaka, H. and the QUEST-RA study group. (2010) Work disability remains a major problem in rheumatoid arthritis in the 2000s: data from 32 countries in the QUEST-RA Study. *Arthritis Research and Therapy* [online]. 12 (2) Available from: http://www.arthritis-research.com/content/12/2/R42 [Accessed 17 December 2014].

Sokka, T., Mäkinen, H., Hannonen, P. and Pincus, T. (2007) Most people over age 50 in the general population do not meet ACR remission criteria or OMERACT minimal disease activity criteria for rheumatoid arthritis. *Rheumatology*. 46 (6), pp. 1020-1023.

Soubrier, M., Zerkak, D., Gossec, L., Ayral, X., Roux, C. and Dougados, M. (2006) Which variables best predict change in rheumatoid arthritis therapy in daily clinical practice? *The Journal of Rheumatology*. 33 (7), pp. 1243-1246.

Spies, C.M., Cutolo, M., Straub, R.H., Burmester, G. and Buttgereit, F. (2010) More night than day — circadian rhythms in polymyalgia rheumatica and ankylosing spondylitis. *The Journal of Rheumatology*. 37 (5), pp. 894-899.

Sprangers, M.A.G. (2010) Disregarding clinical trial-based patient-reported outcomes is unwarranted: Five advances to substantiate the scientific stringency of quality-of-life measurement. *Acta Oncologica*. 49 (2), pp. 155-163.

Sprangers, M.A.G., de Regta, E.B., Andriesb, F., van Agtc, H.M.E., Bijld, R.V., de Boere, J.B., Foetsf, M., Hoeymansg, N., Jacobsh, A.E., Kempeni, G.I.J.M., Miedemaj, H.S., Tijhuisg, M.A.R. and de Haesa, H.C.J.M. (2000) Which chronic conditions are associated with better or poorer quality of life? *Journal of Clinical Epidemiology*. 53, pp. 895–907.

Sprangers, M.A.G. and Schwarz, C.E. (1999) Integrating response shift into health-related quality of life research: a theoretical model. *Social Science and Medicine*. 48, pp. 1507-1515.

Stack, R.J., Sahni, M., Mallen, C.D. and Raza, K. (2013) Symptom complexes at the earliest phase of rheumatoid arthritis: a synthesis of the qualitative literature. *Arthritis Care and Research*. 65 (12), pp. 1916-1926.

Staniszewska, S., Haywood, K.L., Brett, J. and Tutton, L. (2012) Patient and public involvement in patient-reported outcome measures: evolution not revolution. *Patient*. 5 (2), pp. 79-87.

Stanley, K. (2009) Exploring Impact: public involvement in NHS, public health and social care research. *INVOLVE* [online]. Available from: http://www.invo.org.uk/wp-content/uploads/2011/11/Involve_Exploring_Impactfinal28.10.09.pdf [Accessed 28 October 2013].

Strand, V., Holt, R.J., Litman, H.J., Kent, J.D., Pashova, H., Nguyen, J.T. and Etzel, C.J. (2015) Persistence of patient reported morning (AM) stiffness in a large US registry cohort of rheumatoid arthritis (RA) patients initiating new DMARD therapy. *Arthritis and Rheumatology*. 67 (Suppl 10), pp. 3223-3225.

Straub, R.H. and Cutolo, M. (2007) Circadian rhythms in rheumatoid arthritis. Implications for pathophysiology and therapeutic management. *Arthritis and Rheumatism*. 56 (2), pp. 399-408.

Stevens, J.P. (2002) *Applied Multivariate Statistics for the Social Sciences*. London: Lawrence Erlbaum Associates, Publishers.

Streiner, D.L. and Norman, G.R. (2008) *Health Measurement Scales. A Practical Guide to Their Development and Use.* Oxford: Oxford University Press.

Stucki, G., Cieza, A., Geyh, S., Battistella, L., Lloyd, J., Symmons, D., Kostanjsek, N. and Schouten, J. (2004) ICF core sets for rheumatoid arthritis. *Journal of Rehabilitation Medicine*. Suppl 44, pp. 87-93.

Stucki, G., Liang, M.H., Stucki, S., Bruhlmann, P. and Michel, B.A. (1995) A self-administered rheumatoid arthritis disease activity index (RADAI) for epidemiologic research. Psychometric properties and correlation with parameters of disease activity. *Arthritis and Rheumatism*. 38 (6), pp. 795-798.

Stull, D.E., Leidy, N.K., Parasuraman, B. and Chassany, O. (2009) Optimal recall periods for patient-reported outcomes: challenges and potential solutions. *Current Medical Research and Opinion*. 25 (4), pp. 929-942.

Sugiyama, D., Nishimura, K., Tamaki, K., Tsuji, G., Nakazawa, T., Morinobu, A, and Kumagai, S. (2010) Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. *Annals of the Rheumatic Diseases*. 69, pp. 70-81.

Swezey, R.L., Spiegel, T.M., Cretin, S. and Clements, P. (1979) Arthritic hand response to pressure gradient gloves. *Archives of Physical Medicine and Rehabilitation*. 60 (8), pp. 375-377.

Symmons, D.P., Bankhead, C.R., Harrison, B.J., Brennan, P., Barrett, E.M., Scott, D.G.I. and Silman, A.J. (1997) Blood transfusion, smoking, and obesity as risk factors for the development of rheumatoid arthritis. *Arthritis and Rheumatism*. 40 (11), pp. 1955-1961.

Symmons, D., Turner, G., Webb., P., Barrett, A.E., Lunt, M., Scott, D. and Silman, A. (2002) The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology*. 41, pp. 793-800.

Taal, E., Abdel-Nasser, A.M., Rasker, J.J. and Wiegman, O. (1998) A self-report thompson articular index: what does it measure. *Clinical Rheumatology*. 17, pp. 125-129.

Tabachnick, B.G. and Fidell, L.S. (2001) *Using Multivariate Statistics*. Boston: Allyn and Bacon.

Tashakkori, A. and Creswell, J.W. (2007) Editorial: exploring the nature of research questions in mixed methods research. *Journal of Mixed Methods Research*. 1 (3), pp. 207-211.

Tashakkori, A. and Teddlie, C. (1998) *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. London: SAGE Publications Ltd.

Tennant, A. and 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 Care and Research*. 57 (8), pp. 1358-1362.

Terwee, C.B., Bot, S.D.M., de Boer, M.R., van der Windt, D.A.W.M., Knol, D.L., Dekker, J., Bouter, L.M. and de Vet, H.C.W. (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*. 60, pp. 34-42.

Terwee, C.B., Jansma, E.P., Riphagen, I.I. and de Vet, H.C.W. (2009) Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research*. 18, pp. 1115-1123.

Tijhuis, G.J. de Jong, Z., Zwinderman, A.H., Zuijderduin, W.M., Jansen, L.M.A., Hazes, J.M.W. and Vliet Vlieland, T.P.M. (2001) The validation of the Rheumatoid Arthritis Quality of Life (RAQoL) questionnaire. *Rheumatology*. 40, pp. 1112-1119.

Tinsley, H.E. and Tinsley, D.J. (1987) Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*. 34 (4), pp. 414-424.

Tong, A., Sainsbury, P. and Craig, J. (2007) Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*. 19 (6), pp. 349-357.

Tourangeau, R. (1984). Cognitive sciences and survey methods. In: Jabine, T., Straf, M., Tanur, J. and Tourangeau, R. eds. *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Washington, DC: National Academy Press, pp. 73-81.

Tourangeau, R. (2003). Cognitive aspects of survey measurement and mismeasurement. *International Journal of Public Opinion Research*. 15 (1), pp. 3-7.

Tuominen, R., Tuominen, S. and Möttönen, T. (2011) How much is a reduction in morning stiffness worth to patients with rheumatoid arthritis? *Scandinavian Journal of Rheumatology*. 40 (Suppl 125), pp. 12-16.

Tuominen, R., Tuominen, S. and Möttönen, T. (2012) Consistency of assessments and willingness to pay for a reduction in morning symptoms over time in patients with rheumatoid arthritis. *Scandinavian Journal of Rheumatology*. 41 (6), pp. 438-441.

Twohig, H., Mitchell, C., Mallen, C., Adebajo, A. and Mathers, N. (2015a) “I suddenly felt I’d aged”: a qualitative study of patient experiences of polymyalgia rheumatica (PMR). *Patient Education and Counseling*. 98, pp. 645-650.

Twohig, H., Mitchell, C., Mallen, C., Adebajo, A. and Mathers, N. (2015b) Comment on: Stiffness is more than just duration and severity: a qualitative exploration in people with rheumatoid arthritis. *Rheumatology* [online]. Available from: https://rheumatology.oxfordjournals.org/content/54/4/615.short/reply#rheumatology_el_181 [Accessed 6 May 2015].

US Department of Health and Human Services, Food and Drug Administration. (2009) *Guidance for industry: patient-reported outcome measures. Use in medical product development to support labeling claims*. Available from: www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. [Accessed 15 December 2012].

van der Heijde, D.M.F.M., van't Hof, M.A., van Riel, P.LC.M., Theunisse, L.A.M., Lubberts, E.W. van Leeuwen, M.A., van Rijswijk, M.H. and van de Putte, L.B.A. (1990) Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Annals of the Rheumatic Diseases*. 49, pp. 916-920.

van der Heijde, D.F.M., van't Hof, M., van Riel, P.L.C.M. and van de Putte, L.B.A. (1993) Development of a disease activity score based on judgment in clinical practice by rheumatologists. *The Journal of Rheumatology*. 20 (3), pp. 579-581.

van der Heijde, D.M.F.M., van't Hof, M.A., van Riel, P.LC.M., van Leeuwen, M.A., van Rijswijk, M.H. and van de Putte, L.B.A. (1992) Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. *Annals of the Rheumatic Diseases*. 51, pp. 177-181.

van der Linden, M.P.M., le Cessie, S., Raza, K., van der Woude, D., Knevel, R., Huizinga, T.W.J. and van der Helm-van Mil, A.H.M. (2010) Long-term impact of delay in assessment of patients with early arthritis. *Arthritis and Rheumatism*. 62 (12), pp. 3537-3546.

van der Linden, S., Valkenburg, H.A. and Cats, A. (1984) Evaluation of diagnostic criteria for ankylosing spondylitis. *Arthritis and Rheumatism*. 27 (4), pp. 361-368.

van, Nies, J.A.B., Alves, C., Radix-Bloemen, A.L.S., Gaujoux-Viala, C., Huizinga, T.W.J., Hazes, J.M.W., Brouwer, E., Fautrel, B. and Helm-van Mil, A.H.M. (2015) Reappraisal of the diagnostic value of morning stiffness in arthralgia and early arthritis: results from the Groingen EARC, Leiden EARC, ESPOIR, Leiden EAC and REACH. *Arthritis Research and Therapy* [online]. Available from: http://arthritis-research.biomedcentral.com/articles/10.1186/s13075-015-0616-3 [Accessed 6 June 2015].

van Riel, P.L.C.M. and Schumacher, H.R. (2001) How does one assess early rheumatoid arthritis in daily clinical practice? *Best Practice and Research Clinical Rheumatology*. 15, pp. 67-76.

van Riel, P.L.C.M. and van de Putte, L.B.A. (1994) Clinical assessment and clinical trials in rheumatoid arthritis. *Current Opinion in Rheumatology*. 6, pp. 132-139.

van Riel, P.L.C.M. and van Gestel, A.M. (2000) Clinical outcome measures in rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 59 (Suppl 1), pp. i28-i31.

van Tubergen, Debats, I., Ryser, L., Londono, J., Burgos-Vargas, R., Cardie, M.H., Landwew, R., Stucki, G. and van der Heijde, D. (2002) Use of a numerical rating scale as an answer modality in ankylosing spondylitis–specific questionnaires. *Arthritis Care and research*. 47 (3), pp. 242-248.

van Tuyl, L.H.D., Britsemmer, K., Wells, G.A., Smolen, J.S., Zhang, B., Funovits, J., van Schaardenburg, D., Felson, D. and Boers, M. (2011) Remission in early rheumatoid arthritis defined by 28 joint counts: limited consequences of residual disease activity in the forefeet on outcome. *Annals of the Rheumatic Diseases*. 71, pp. 33-37.

van Tuyl, L.H.D., Hewlett, S., Sadlonova, M. Davis, B., Flurey, C., Hoogland, W., Kirwan, J., Sanderson, T., van Schaardenburg, D., Scholte-Voshaar, M., Smolen, J., Stamm, T. and Boers, M. (2015) The patient perspective on remission in rheumatoid arthritis: 'You've got limits but you're back to being you again'. *Annals of the Rheumatic Diseases*. 74, pp. 1004-1010.

van Tuyl, L.H.D., Lems, W.F. and Boers, M. (2014) Measurement of stiffness in patients with rheumatoid arthritis in low disease activity or remission: a systematic review. *BMC Musculoskeletal Disorders* [online]. Available from: http://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/1471-2474-15-28 [Accessed 15 March 2014].

Velicer, W.F., Eaton, C.A. and Fava, J.L. (2000) Construct explication through factor or component analysis: a review and evaluation of alternative procedures for determining the number of factors or components. In Goffin, R.D. and Helmes, E. ed. (2000) *Problems and Solutions in Human Assessment*. London: Kluwer Academic Publishers, pp. 41-71.

Verstappen, S.M.M., Jacobs, J.W.G. and Hyrich, K.L. (2007) Effect of anti-tumor necrosis factor on work disability. *The Journal of Rheumatology*. 34 (11), pp. 2126-2128.

Vliet Vlieland, T.P.M., Zwinderman, A.H., Breedveld, F.C. and Hazes, J.M.W. (1997) Measurement of morning stiffness in rheumatoid arthritis clinical trials. *Journal of Clinical Epidemiology*. 50 (7), pp. 757-763.

von Thurn, D. and Moore, J. (1994) Results from a cognitive exploration of the 1993 American housing survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association [online]. Available from: https://www.census.gov/srd/papers/pdf/sm9403.pdf [Accessed 22 June 2015].*

Ward, M.M. (1994) Clinical measures in rheumatoid arthritis: which are most useful in assessing patients. *The Journal of Rheumatology*. 21 (1), pp. 17-27.

Ward, M.M., Guthrie, L.C. and Alba, M. (2015) Measures of arthritis activity associated with patient-reported improvement in rheumatoid arthritis when assessed prospectively versus retrospectively. *Arthritis Care and Research*. 67 (6), pp. 776-871.

Wellens, T. (1994) The cognitive evaluation of the nativity questions for the current population survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association [online].* Available from: http://www.amstat.org/sections/srms/Proceedings/ [Accessed 22 June 2015].

Wells, G., Boers, M., Shea, B., Brooks, P.M., Simon, Strand, V., Aletaha, D., Anderson, J.J., Bombardier, C., Dougados, M., Emery, P., Felson, D.T., Fransen, J., Furst, D.E., Hazes, J.M.W., Johnson, K.R., Kirwan, J.R., Landewé, R.B.M., Lassere, M.N.D., Michaud, K., Suarez-Almazor, M., Silman, A.J., Smolen, J.S., van der Heijde, D.M.F.M., van Riel, P.L.C.M., Wolfe, F. and Tugwell, P.S. (2005) Minimal disease activity for rheumatoid arthritis: a preliminary definition. *The Journal of Rheumatology*. 32 (10), pp. 2016-2024.

Westhoff, G., Buttgereit, F., Gromnica-Ihl, E. and Zink, A. (2008) Morning stiffness and its influence on early retirement in patients with recent onset rheumatoid arthritis. *The Journal of Rheumatology*. 47, pp. 980-984.

White, J. and Bryer, D. (2006) Medications in Rheumatic Diseases. In: Hill, A. ed. (2006) *Rheumatology Nursing. A Creative Approach.* Chichester: John Wiley & Sons Ltd, pp. 337-372.

White, D.K., Wilson, J.C. and Keysor, J.J. (2011) Measures of adult general functional status. *Arthritis Care and Research*. 63 (Suppl 11), pp. s297-s307.

Wiesinger, T., Smolen, J.S., Aletaha, D. and Stamm, T. (2013) Compression test (gaenslen's squeeze test) positivity, joint tenderness, and disease activity in patients with rheumatoid arthritis. *Arthritis Care and Research*. 65 (4), pp. 653-657.

Williamson, A. and Hoggart, B. (2005) Pain: a review of three commonly used pain rating scales. *Issues in Clinical Nursing*. 14 (7), pp. 798-804.

Willis, G.B. (1999) Cognitive interviewing: a "how to" guide. *Research Triangle Institute* [online]. Available from: http://appliedresearch.cancer.gov/archive/cognitive/interview.pdf#search=willis+how+to+guide [Accessed 16 May 2014].

Willis, G.B. (2005) *Cognitive Interviewing. A Tool for Improving Questionnaire Design*. London: Sage Publications Ltd.

Willis, G.B. and Lessler, J.T. (1999) Question appraisal system (QAS). *Research Triangle Institute* [online]. Available from: https://www2a.cdc.gov/prams/pubs_references/Questionnaires/QAS_99.PDF [Accessed 23 May 2014].

Wilson, O. (2016) *The impact, frequency and care of foot problems in people with rheumatoid arthritis*. PhD, University of the West of England.

Wilson, O., Hewlett, S., Pollock, J., Woodburn, J., Quest, E., Swales, C. and Kirwan, J. (2015) Population survey of the prevalence, impact and care of foot problems in people with rheumatoid arthritis. *Journal of Foot and Ankle Research*. 8 (Suppl 1), pp. A9.

Wolfe, F. (1997) Comparative usefulness of C-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. *The Journal of Rheumatology*. 24 (8), pp. 1477-1485.

Wolfe, F. (1999) Determinants of WOMAC function, pain and stiffness scores: evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. *Rheumatology*. 38, pp. 355-361.

Wolfe, F., Allaire, S. and Michaud, K. (2007) The prevalence and incidence of work disability in rheumatoid arthritis, and the effect of anti-tumor necrosis factor on work disability. *The Journal of Rheumatology*. 34 (1), pp. 2211-2217.

Wolfe, F., Hawley, D.J. and Wilson, K. (1996) The prevalence and meaning of fatigue in rheumatic disease. *The Journal of Rheumatology*. 23 (8), pp. 1407-1417.

Wolfe, F., Michaud, K. and Pincus, T. (2005) A composite disease activity scale for clinical practice, observational studies, and clinical trials: the patient activity scale (PAS/PAS-II). *The Journal of Rheumatology*. 32 (12), pp. 2410-2415.

World Health Organisation. (2002) *Towards a common language for functioning, disability and health. ICF. The international classification of functioning, disability and health.* Available from: http://www.who.int/classifications/icf/en/ [Accessed 21 October 2015].

World Health Organisation. (2015) *Disabilities*. Available from: http://www.who.int/topics/disabilities/en/ [Accessed 15 June 2013].

Wright, V., Dowson, D. and Longfield, M.D. (1969) Joint stiffness-its characterisation and significance. *Biomedical Engineering*. 4 (1), pp. 8-14.

Wright, V. and Johns, R.J. (1960) Observations on the measurement of joint stiffness. *Arthritis and Rheumatism*. 3 (4), pp. 328-340.

Wright, V. and Johns, R.J. (1961) Quantitative and qualitative analysis of joint stiffness in normal subjects and in patients with connective tissue disease. *Annals of the Rheumatic Diseases*. 20, pp. 36-46.

Wylde, V., Wells, V., Dixon, S. and Gooberman-Hill, R. (2014) The colour of pain: can patients use colour to describe osteoarthritis pain? *Musculoskeletal Care*. 12 (1), pp. 34-46.

Yazici, Y., Pincus, T., Kautiainen, H. and Sokka, T. (2004) Morning stiffness in patients with early rheumatoid arthritis is associated more strongly with functional disability than with joint swelling and erythrocyte sedimentation rate. *The Journal of Rheumatology*. 31 (9), pp. 1723-1726.

Yelin, E., Henke, C. and Epstein, W. (1987) The work dynamics of the person with rheumatoid arthritis. *Arthritis and Rheumatism*. 30 (5), pp. 507-512.

Youngson, A., Cole, F., Wilby. And Cox, D. (2015) The lived experience of diabetes: conceptualization using a metaphor. *The British Journal of Occupational Therapy*. 78 (1), pp. 24-32.

Zakeri, Z., Sandoughi, M., Mashhadi, M.A., Raesi, V. and Shahbakhsh, S. (2013) Serum vitamin D level and disease activity in patients with recent onset rheumatoid arthritis. *International Journal of Rheumatic Diseases* [online]. Available from: http://onlinelibrary.wiley.com/doi/10.1111/1756-185X.12181/epdf [Accessed 28 December 2014].

Appendices

Appendix A: Data extraction form

Stiffness measurement tools – data screening sheet

Review details

Reviewer name	
Date performed	
Study title, authors, date	

Eligibility

		Yes/No/?	Notes/Evidence
Does the paper include populations with RA?*			
Does the paper report on measurement properties of stiffness?*			
How does the study report on stiffness:	Stiffness as an outcome in relation to other core set disease activity measures+		
	The development of a patient reported tool to measure stiffness+		
	A comparison of two or more different tools to measure aspects of stiffness+		
Include the study in the review?			

*Round 2 screening from van Tuyl (2014); +Round 3 screening from van Tuyl (2014)

Comments/queries

Stiffness measurement tools – data extraction sheet

Study description

Study design	
Study objectives	
Study participants (age, gender, diagnosis)	

Stiffness items

Number of items identified	Item concept	Item wording	Item response options or anchors	Item timeframe
1				
Add more if required				

For each of the below criteria please report:

- Whether each was reported on in the paper or not
- If it was reported please detail how (including: what was compared, what tests were used, what the results were, what population/s this was performed in, and any other relevant information)
- Complete one report for every item; if there is more than 1 item per paper, please report for each item individually
- In circumstances where studies refer to other papers to describe the item development (e.g. particularly for content validity), or validation (e.g. particularly for construct validity) process. State what information is provided by the original paper and then perform a separate review on the referenced paper. All information for each instrument can be collated later.

Validity

Face validity (credibility; does the instrument look sensible?)	
Consider: i. The concept being assessed, on what anchors ii. The length and complexity of the tool iii. Whether stiffness is self-reported	

Content validity (comprehensiveness; does the instrument contain all relevant content that is important to the intended populations?)	
Consider: i. Qualitative evidence that items and domains are relevant to the intended population, the measurement concept, and for the tools intended use	

Criterion validity (accuracy; does the instrument perform well against a gold standard?)	
Consider: i. How the concept was assessed in the different methods ii. Are the methods correlated? iii. Do the methods perform equally well (e.g. sensitivity)	

Validity continued

Construct validity (biological sense; do results agree with expected hypotheses?)	
Convergent validity (does the instrument demonstrate relationships where they would be expected?) Consider: i. The hypotheses tested ii. Correlations	
Divergent validity (does the instrument demonstrate no relationship when no relationship is expected?) Consider: i. The hypotheses tested ii. Correlations	
Discriminant/known groups validity (does the instrument demonstrate ability to differentiate between expected groups? E.g. inflam (RA) and non-inflam (OA)) Consider: i. The hypotheses tested ii. Correlations	

Reliability

Test-retest reliability (reproducibility)	
Consider: i. Time interval	

Internal consistency

Internal consistency (agreement among items in subscales)	
Consider: ii. Cronbach's alpha	

Ability to detect change

Ability to detect change (responsiveness; when patient experiences change, does the instrument score reflect that change?)	
Consider: i. Is change in the instrument score also seen in similar measures that indicate the patient state has changed? ii. Sensitivity iii. Full range iv. Responsiveness to therapy?	

Floor and ceiling effects

Floor and ceiling effects (can the instrument distinguish responders with lowest or highest possible score?)	
Consider: i. Do more than 15% responders achieve lowest or highest possible score?	

Interpretability

Interpretability (can the instrument provide information about what change in score would be clinically meaningful?)	
Consider: i. Means and SDs ii. Population norms	

Appendix B: Patient information sheet (Study 1)

A research study to explore stiffness in people with rheumatoid arthritis (Phase 1)

Patient information sheet

You are being invited to take part in a research study. Before you decide whether to take part, it is important for you to understand why the research is being done, and what it will involve. Please take time to read the following information leaflet carefully and discuss it with others if you wish. If anything is not clear or you would like more information please ask one of the team.

What is the purpose of the study?
One of the problems commonly experienced by people with rheumatoid arthritis is stiffness. Stiffness is a term used by patients and health professionals, but it is not very well understood. The aim of this study is to understand your experience of stiffness and the language that you use to describe it. Your knowledge will help improve the current understanding of stiffness. It will also help to develop a way of measuring stiffness for people with rheumatoid arthritis.

Why have I been chosen?
You have been chosen because you have rheumatoid arthritis.

Who is asking me to take part?
I am Serena Halls, a PhD student at the University of the West of England. This research study is the first of three studies which will form part of my PhD.

What will I be asked to do if I take part?
You will be asked to attend an informal one to one interview with the researcher (Serena Halls) at the rheumatology department you attend. Before the interview starts, the researcher will ask you to read and sign the consent form and ask you some questions about your medical history. In the interview, she will invite you to discuss your experience of stiffness. You can say as much or as little as you like, there are no right or wrong answers - we are looking for your own individual experience. The interview will last for about an hour and we will offer you refreshments and are happy to pay your travel costs. We will ask your permission to audio-record the interview, which we will type up (transcribe) and then analyse after the interview.

Do I have to take part?
No, taking part is voluntary. It is up to you to decide whether or not to take part. If you decide not to take part you do not need to give a reason, nobody will be upset and the standard of care you receive will not be affected. If you decide to take part we will ask you to sign a consent form, and will give you a copy of this information sheet and the consent form to keep.

For general advice about taking part in research, you can contact the local Patient Advice and Liaison Service on 0117 900 3433 or pals@bristolpct.nhs.uk.

What if I wish to withdraw at a later stage?
You are free to withdraw from the study at any time, and with no explanation.

What are the risks of taking part in the study?
We do not believe there are any risks in being involved in this study. We appreciate that there may be some inconvenience to you by having to come into the hospital for the interview but we will try and reduce this by arranging a convenient date and time for you to come.

What are the benefits of taking part in the study?
The benefits of taking part in this study are that you will be helping us to gain a better understanding of stiffness in people with rheumatoid arthritis. This will help us to improve decisions made about treatment and management of rheumatoid arthritis.

Will my taking part in the study be kept confidential?
Yes. Your name will be replaced by a code. All other identifying information (such as people's names, locations or specific descriptions) will be replaced with code numbers or a generalised summary. No one will be able to identify you from any analysis or report. The study reports will include quotations from the interviews but no names will be used. The recordings will be kept securely for 6 years and then destroyed, in accordance with best practice in research guidelines.

What will happen to the results of this research study?
Research team members will analyse the anonymous transcripts and discuss our findings. The findings of this study will influence the design of later research studies within this PhD. We hope the results will be reported in professional journals and at meetings (but participants will not be identified by name). We will send you a summary of the results if you would like.

Who is organising, funding and reviewing the research?
The study is coordinated by a team from the university of the West of England (UWE) based at the Academic Rheumatology Unit at the Bristol Royal Infirmary. It is funded by UWE and has been peer reviewed by the local and UWE Research Ethics Committees.

What do I do now?
Thank you for considering taking part in this research. Please return the reply slip provided if you would like to take part by returning it in the pre-paid reply envelope to Serena Halls. Serena will then contact you in a few days with further information and to answer any questions.

Research team:

Serena Halls, PhD Student Researcher, UWE Bristol
Professor Sarah Hewlett, Professor of Rheumatology Nursing, UWE Bristol
Professor John Kirwan, Professor of Rheumatic Diseases, UoB Bristol
Dr Jon Pollock, Reader in Epidemiology, UWE Bristol
Dr Emma Dures, Research Fellow, UWE Bristol
Mrs Avis Edmunds, Patient Research Partner
Mrs Gill Baker, Patient Research Partner

Contact:

Serena Halls	0117 342 4972	Serena.Halls@uwe.ac.uk
Sarah Hewlett	0117 324 2903	Sarah.Hewlett@uwe.ac.uk

Appendix C: Interview topic guide with prompts (Study 1)

A. Can you tell me about your experience of stiffness in relation to RA?
    a. Definition?
    b. Sensation? Is there anything that feels similar to the feeling of stiffness?
    c. Is stiffness the right word to describe how you feel?
    d. If you were to describe stiffness to someone who didn't have rheumatoid arthritis and didn't really know what it was, what would you say?
    e. Where in your body do you feel stiff? Is it always the same place? Does it feel the same?

B. How does this vary in a 24 hour period?
    a. Does it gradually go away or is it sudden?
    b. Is it different during the day or night?
    c. Causes?
    d. Differences at different times of disease activity?

C. Has stiffness changed over the course of your disease?
    a. Before you had rheumatoid arthritis did you ever experience that same feeling?
    b. Before you had rheumatoid arthritis did you experience a different feeling that you would call/class as stiffness?
    c. Do you have any other diseases that make you stiff?

D. How does stiffness differ from other RA symptoms?
    a. Pain/fatigue/other
    b. What is the relationship between symptoms?
    c. How relevant is stiffness in relation to other symptoms?

E. What are the consequences of stiffness?

F. How do you deal with stiffness?
    a. Influence of medications?

G. How to you assess stiffness?
    a. When you are asked in clinic by a consultant or nurse about stiffness, what do you say?
    b. What are your thoughts about stiffness measurement?
    c. How do you know when it is good or bad?

H. Is there anything that you feel is important to stiffness that we have not talked about?

I. If your stiffness was an animal what would it be and why?

Appendix D: Questionnaire pack

Stiffness pre-discussion questionnaire

Date: _____ **Study ID:** _____

Thank you for agreeing to take part in this study

This questionnaire will help the researchers make sure that they talk to a wide range of people who experience stiffness. Your answers are confidential to the researchers, and although other people will see the results of the overall study, they will not be able to link your name to the answers you give on this sheet.

A) This section asks about your socio demographic details

1. Gender: Male / Female (Please circle)

2. Date of birth: __ __ **/** __ __ **/** __ __ __ __ (Day/Month/Year)

3. Work status: Paid work (Please circle)
   Student
   Homemaker
   Unemployed
   Retired
   Receiving incapacity benefits
   Other: _____ (Please specify)

4. Education: Did not complete school (Circle highest level)
   School education
   College / apprenticeship
   University level education
   Other: _____ (Please specify)

B) This section asks about your rheumatoid arthritis

1. How long have you been diagnosed with rheumatoid arthritis?

   _____ (Years)

2. Have you ever experienced stiffness related to your rheumatoid arthritis?

   Yes No (Please circle)

3. Please list the medications you are taking for your rheumatoid arthritis:

_____

_____

_____

_____

_____

4. Have you started or changed medication in the last six weeks?

Yes No (Please circle)

5. Have you had a steroid injection in the past two months?

Yes No (Please circle)

6. How much pain have you had because of your arthritis in the PAST WEEK? (Place a vertical line to indicate the severity of the pain)

No pain |_____| Severe pain

7. Considering all of the ways your arthritis affects you, please mark on the line to show how well you are doing:

Very well |_____| Very badly

C) This section asks about your usual ABILITIES over the PAST WEEK (Please tick)

	Without ANY difficulty	**With SOME difficulty**	**With MUCH difficulty**	**Unable to do**
1. DRESSING AND GROOMING Are you able to:				
- Dress yourself, including tying shoelaces and doing buttons?	_____	_____	_____	_____
- Shampoo your hair?	_____	_____	_____	_____
2. RISING Are you able to:				
- Stand up from an armless straight chair?	_____	_____	_____	_____
-Get in and out of bed?	_____	_____	_____	_____
3. EATING Are you able to:				
- Cut your meat?	_____	_____	_____	_____
- Lift a full cup or glass to your mouth?	_____	_____	_____	_____
- Open a new carton of milk (or soap powder)?	_____	_____	_____	_____
4. WALKING Are you able to:				
- Walk outdoors on flat ground?	_____	_____	_____	_____
- Climb up five steps?	_____	_____	_____	_____

Please tick any aids or devices that you usually use for any of these activities:

_____ Cane
_____ Walking frame
_____ Crutches
_____ Wheelchair
_____ Devices used for dressing (button hook, zipper pull, long handled shoe horn etc)
_____ Built-up or special utensils
_____ Special or built-up chair

Other: _____ (Please specify)

Please tick any categories for which you usually need help from another person:

_____ Dressing and grooming
_____ Rising
_____ Eating
_____ Walking

	Without ANY difficulty	With SOME difficulty	With MUCH difficulty	Unable to do
5. HYGIENE Are you able to: - Wash and dry your entire body?				
- Take a bath?				
- Get on and off the toilet?				
6. REACH Are you able to: - Reach and get down a 5lb object (e.g. a bag of potatoes) from just above your head?				
- Bend down to pick up clothing from the floor?				
7. GRIP Are you able to: - Open car doors?				
- Open jars which have been previously opened?				
- Turn taps on and off?				
8. ACTIVITIES Are you able to: - Run errands and shop?				
- Get in and out of a car?				
- Do chores such as vacuuming, housework or light gardening?				

Please tick any aids or devices that you usually use for any of these activities:

_____ Raised toilet seat _____ Bath rail
_____ Bath seat _____ Long handled appliances for reach
_____ Jar opener (for jars previously opened)

Other: _____ (Please specify)

Please tick any categories for which you usually need help from another person:

_____ Hygiene _____ Gripping and opening things
_____ Reach _____ Errands and housework

THANK YOU FOR COMPLETING THIS QUESTIONNAIRE

Appendix E: Study 1 coding tree

Codes	Level 1 groups	Level 2 groups	Theme
Stiffness is normal	Stiffness is normal	Stiffness is a normal consequence of RA	Part of my disease
No stiffness prior to RA	No stiffness prior to RA		
Stiffness is an obvious symptom of RA	Stiffness is an obvious symptom of RA		
Stiffness is part of a bigger picture in the morning specifically			
Stiffness is part of a bigger picture			
RA stiffness is different – General	RA stiffness is specific		
RA stiffness is different to exercise stiffness			
RA stiffness is different to overuse stiffness			
RA stiffness is different to OA stiffness			
RA stiffness is not different to OA stiffness			
During a period of flare you cannot move	Stiffness relates to flare	Stiffness varies with disease fluctuations	
During a period of flare it is difficult to move and you can't do ADLs			
Can't do anything in a flare up			
Stiffness affects sleep when it's bad			
During a period of flare pain and stiffness more related			
During a period of flare stiffness and inflammation more related			
During a period of flare all symptoms are worse			
During a period of flare stiffness does not go away			
During a period of flare stiffness sticks around			
During a period of flare stiffness is quick/sudden			
In flare stiffness lasts longer			
In flare stiffness is more frequent in occurrence			
During a period of flare stiffness affects more joints			
During a period of flare stiffness severity is high			
During a period of flare stiffness is an exaggeration of itself			
During a period of flare stiffness is acute			

Codes	Level 1 groups	Level 2 groups	Theme
During a period of flare stiffness you can't use manage it in the same way			
Harder to ease in flare			
Damaged joints cause stiffness	Stiffness relates to damaged joints		
You can't move damaged joints			
Causes of damaged joints			
Stiffness from damage and flare are different			
Pain and stiffness are normal	Patient symptoms – Pain	Relationship between stiffness and other RA symptoms	
Pain and stiffness are related			
Pain and stiffness are different concepts			
Pain comes with baggage			
Pain is easier to deal with than stiffness			
Pain is harder to deal with than stiffness			
Pain and stiffness more related during flare			
Pain effects sleep			
You feel stiff when you try to move but pain you feel when you don't move			
If you stay in one position for too long you get pain			
You can tell pain and stiffness limitations apart			
Stiffness and fatigue are related	Patient symptoms - Fatigue		
Stiffness is more fatigue than pain			
Stiffness and fatigue are different			
Stiffness and fatigue are distinguishable by the time they occur			
Stiffness is related to inflammation	Medical symptoms – Inflammation		
Stiffness is unrelated to inflammation			
Inflammation causes stiffness			
Inflammation relates to a flare			
Stiffness is the same as inflammation			

Codes	Level 1 groups	Level 2 groups	Theme
Both stiffness and inflammation cause functional loss			
Patients don't recognise inflammation			
All symptoms are intertwined	All symptoms are intertwined		
All symptoms are intertwined when its bad but unrelated when it's not			
All symptoms are worse [in a flare]			
Symptoms are not just related when disease is bad			
Stiffness was an early symptom of RA	Stiffness relevant in early RA	Varying prominence of stiffness during course of the disease	
Stiffness was severe at disease onset			
Stiffness was not relevant in early disease	Stiffness not relevant in early RA		
Stiffness more relevant later in disease duration			
Stiffness affects certain parts of the body	Stiffness affects certain locations	Location within body	Local and widespread
Stiffness feels the same in different part of the body			
Stiffness moves around the body	Stiffness affects variable locations		
Stiffness feels different in different parts of the body			
During a flare/bad day stiffness affects more joints	Stiffness all over		
In the morning stiffness affects more joints/all over			
Joints are the problem	Stiffness affects particular structures	Affected body structures	
Muscles are the problem			
Tendons are the problem			
Uncertainty about what structure is the problem	Uncertainty about affected structures		
Stiffness is caused by being immobile	Immobility	Movement and stiffness	Linked to behaviour and environment
Being in a restricted position causes stiffness			
Should be less stiff when have been more active in the night but not the case			
Legacy of activity causes stiffness	Overdoing it		
Overdoing it and not resting causes stiffness			
Medications are beneficial for stiffness – Non-specific			

Codes	Level 1 groups	Level 2 groups	Theme
Medications are beneficial for stiffness – Infusion	Medications have an impact on stiffness	Medications and stiffness	
Medications are beneficial for stiffness – Steroids			
Medications are beneficial for stiffness – DMARD			
Medications are beneficial for stiffness - Anti-TNF			
Medications reduce stiffness duration	Medications reduce stiffness		
Medications reduce stiffness severity			
Medications reduce stiffness which allows normality			
Medications get rid of stiffness			
Only steroids reduce stiffness in flare			
Medications reduce stiffness but it's a lost entity because it's never measured	Medications impact on stiffness is lost		
Medications do not target stiffness	Medications are not beneficial for stiffness		
Medications do not work when you have a flare up			
Some medications work better than others	Medications effects can be variable		
Sometimes the same medications work better than other times			
You have transition periods between medications			
Medications have side effects	Medications have other considerations		
Not wanting to take medications			
You don't have a choice but to take medications			
Certain drinks cause stiffness to be worse	Diet	Lifestyle and environment and stiffness	
Certain foods cause stiffness to be worse			
Air conditioning causes joints to be stiff and painful	Weather		
The weather has an impact on stiffness – Hot weather			
The weather has an impact on stiffness – Cold weather			
The weather has an impact on stiffness – Humid weather			
Stiffness does not affect sleep	Sleep		
Stiffness does affect sleep			

Codes	Level 1 groups	Level 2 groups	Theme
Poor sleep impacts on stiffness			
Stiffness is individual/subjective	Individual experience	Stiffness is individual	Highly variable
Stiffness is different for different people	Different experience		
Stiffness does relate to mornings	Morning	Temporal pattern of stiffness	
Stiffness does not relate to mornings			
You notice stiffness when you get up/wake up			
Stiffness relates to other times of day	Other times of day		
Stiffness doesn't relate to a particular time of day			
Best time of day is afternoon/evening			
Timing of stiffness has changed through the course of the disease	Timing changes		
Duration of stiffness – Varies within people	Varies	Duration of stiffness	
Duration of stiffness – Varies between people			
Duration of stiffness - Different amount of time			
Duration of stiffness - Gradually eases off			
Duration of stiffness - Similar amount of time	Constant		
Stiffness never completely goes away			
Stiffness is constant now (it doesn't change)			
Some joints stick all day			
Time is not a relevant factor	Time is not relevant		
General stiffness – Expected	General stiffness	Severity of stiffness	
General stiffness – Impact on function			
General stiffness – Manageable			
Severe stiffness – Cramp stiffness	Severe stiffness		
Severe stiffness – Not expected [harder to manage]			
Can't function when its bad			
Damage related stiffness – Element of severity in damaged joints	Damage related stiffness		
Can't do ADL's or simple tasks		Daily life impact	Impacts on daily life
Restricts hobbies			

Codes	Level 1 groups	Level 2 groups	Theme
Can’t eat	Impact on activities of daily living and essential tasks		
Restricts tasks			
Driving			
Dressing			
Causes difficulties			
Getting up from a chair is difficult			
Stairs			
Restricts family life			
Cant plan			
Difficult to get out of bed			
Unnatural/awkward walking			
Can’t get comfortable			
Stiffness a problem at night			
Stiffness not a problem at night			
Impact on life			
Restriction			
Disabling			
Affects work	Impact on work		
Loose normality	Impact on normality		
Stiffness in different places is worse because it has different impact	Impact differs in different locations		
Location of stiffness affects the impact [worse in hands]			
Certain movements			
Difficult to move	General impact	Physical impact	
Unable to move			
No dexterity	Specific impact		
Reduced ROM			
Grip			
Bending down			
No balance			
Don’t have mobility			

Codes	Level 1 groups	Level 2 groups	Theme
Slow/can't rush/actions take me longer			
Physically have to unbend limbs			
No flexibility			
Need physical support			
Forget how to walk properly	Cognitive impact	Cognitive impact	
Have to think about doing actions			
Psych impact/general wellbeing	Psychological impact	Psychological impact	
Mood affects symptoms			
Stiffness doesn't affect mood			
Stiffness makes you frustrated			
Stress and anxiety impacts on symptoms			
Restricts image – vanity			
You lose the good part of the day			
Pain on movement	Pain impact	Pain impact	
If you force movements you get pain			
Pain on movement is accentuated in flare			
Must move to relieve stiffness which means you go through the pain barrier			
Moving	General moving	Movement based strategies	Requires self-management
Moving position			
Moving while still in bed			
Walking	Specific moving		
Stretching			
Supporting joints			
Physically manipulating your joints			
Balancing rest and movement			
Exercise	Exercise		
Gadgets/aids	Gadgets and splints	Other strategies	
Splints are effective for stiffness			
Splints are not effective for stiffness			

Codes	Level 1 groups	Level 2 groups	Theme
Splints are not compatible with other self-management strategies			
Hot/cold therapy is effective	Heat and cold		
Shower in the morning			
Hot/cold therapy is not effective			
Hydrotherapy			
Alternative therapy - Changing diet	Alternative therapies		
Alternative therapy - Aromatherapy			
Alternative therapy - Relaxation techniques			
Take medication in the morning to get going	Medications and painkillers		
Take medication to function			
Take painkillers			
Family and friends – Physically help with jobs/housework	External – Social support	Psychosocial strategies	
Family and friends - Facilitate use of self-management strategies			
Family and friends - Are flexible			
Normalise/accept stiffness	Internal – Normalise		
Adapt and adjust behaviours generally - Have to find other ways to do things	Internal – Adapt and adjust		
Adapt and adjust behaviours generally - You can work round stiffness			
Adapt and adjust behaviours in the morning - You just adapt in the morning			
Adapt and adjust behaviours in the morning - Perform activities later in day			
Adapt and adjust behaviours in the morning - Use gadgets at difficult times			
Prepare/plan for stiffness - Compensate for slow movement by getting up earlier	Internal – Prepare and plan		
Prepare/plan for stiffness - Not too restrictive if pace and plan			
Self-management is easier since stopping work			
Ensure RA and stiffness does not take over life [RA general]			

Codes	Level 1 groups	Level 2 groups	Theme
Self-management is individual/knowing your limitations/your body [general RA]	Internal – Part of general RA management		
Do what you can to help yourself [general RA]			
Self-management and understanding RA develops over time [RA general]			
Self-management can be difficult especially when you enjoy things			

Appendix F: COREQ checklist (Study 1)

No. Item	Guide questions/description	How was this component addressed
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Interviewer	Which author/s conducted the interview or focus group?	Identified in Section 4.3.4.
2. Credentials	What were the researcher's credentials?	The characteristics of the supervisory team were identified in Section 3.4.2 and the researchers (Halls) background was highlighted in Section 3.2.1
3. Occupation	What was their occupation at the time of the study?	
4. Gender	Was the researcher male or female?	
5. Experience and training	What experience or training did the researcher have?	
Relationship with participants		
6. Relationship established	Was a relationship established prior to study commencement?	
7. Participant knowledge of the interviewer	What did the participants know about the researcher?	Identified in Section 4.3.4.
8. Interviewer characteristics	What characteristics were reported about the interviewer?	The researchers (Halls) background, experience and research interests were highlighted in Section 3.2.1
Domain 2: study design		
Theoretical framework		
9. Methodological orientation and theory	What methodological orientation was stated to underpin the study?	The methodological approach to the study was discussed in Section 3.3.1 and the method and analysis approaches were discussed in Sections 4.3.1 and 4.4.1
Participant selection		
10. Sampling	How were participants selected?	
11. Method of approach	How were participants approached?	Identified in Section 4.3.2 and 4.5.1
12. Sample size	How many participants were in the study?	
13. Non-participation	How many people refused to participate or dropped out? Reasons?	Identified in Section 4.3.4
Setting		
14. Setting of data collection	Where was the data collected?	Identified in Section 4.3.4

No. Item	Guide questions/description	How was this component addressed
15. Presence of non-participants	Was anyone else present besides the participants and researchers?	Identified in Section 4.3.4 and 4.5.1
16. Description of sample	What are the important characteristics of the sample?	
Data collection		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Identified in Section 4.3.3 and Appendix C
18. Repeat interviews	Were repeat interviews carried out? If yes, how many?	This was not performed
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	Identified in Section 4.3.4
20. Field notes	Were field notes made during and/or after the interview or focus group?	Identified in Section 4.3.4
21. Duration	What was the duration of the interviews or focus group?	Identified in Section 4.3.4
22. Data saturation	Was data saturation discussed?	Identified in Section 4.3.2
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	This was not performed as part of Study 1 however, the findings were validated in Study 2
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	Identified in Section 4.4.1.1.2
25. Description of the coding tree	Did authors provide a description of the coding tree?	Identified in Section 4.5.2 and Appendix E
26. Derivation of themes	Were themes identified in advance or derived from the data?	Identified in Sections 4.4.1.1 and 4.5.2
27. Software	What software, if applicable, was used to manage the data?	Identified in Sections 4.4.1.1.2
28. Participant checking	Did participants provide feedback on the findings?	This was not conducted as part of Study 1 however, the findings were validated in Study 2
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes? Was each quotation identified?	Identified in Section 4.5.2
30. Data and findings consistent	Was there consistency between the data presented and the findings?	
31. Clarity of major themes	Were major themes clearly presented in the findings?	
32. Clarity of minor themes	Is there a description of diverse cases or minor themes?	

Appendix G: Patient information sheet (Study 2)

A research study to develop a questionnaire for stiffness in rheumatoid arthritis (Phase 2)

Patient information sheet

You are being invited to take part in a research study. Before you decide whether to take part, it is important for you to understand why the research is being done, and what it will involve. Please take time to read the following information leaflet carefully and discuss it with others if you wish. If anything is not clear or you would like more information please ask one of the team.

What is the purpose of the study?
One of the problems commonly experienced by people with rheumatoid arthritis is stiffness. Stiffness is a term used by patients and health professionals, but it is not very well understood. The aim of this study is to develop a way of measuring stiffness in rheumatoid arthritis. With the help of other people like yourself we have developed a better understanding of what stiffness means to people. We would now like your help to develop this further by working towards a way of measuring stiffness in a questionnaire.

Why have I been chosen?
You have been chosen because you have rheumatoid arthritis.

Who is asking me to take part?
I am Serena Halls, a PhD student at the University of the West of England. This research study is the second of three studies which will form part of my PhD.

What will I be asked to do if I take part?
You will be asked to attend a group discussion with the researcher (Serena Halls) and between 4 and 6 other people like yourself at the rheumatology department you attend. Before the focus group starts, the researcher will ask you to read and sign the consent form and ask you some questions about your medical history. In the group discussion, she will invite you to discuss your thoughts about stiffness and how to measure it. You can say as much or as little as you like, there are no right or wrong answers. The discussion will last for about an hour and we will offer you refreshments and are happy to pay your travel costs. We will ask your permission to audio-record the discussion, which we will type up (transcribe) and then analyse after the discussion finishes.

Do I have to take part?
No, taking part is voluntary. It is up to you to decide whether or not to take part. If you decide not to take part you do not need to give a reason, nobody will be upset and the standard of care you receive will not be affected. If you decide to take part we will ask you to sign a consent form, and will give you a copy of this information sheet and the consent form to keep. For general advice about taking part in research, you can contact the local Patient Advice and Liaison Service on 0117 900 3433 or pals@bristolpct.nhs.uk.

What if I wish to withdraw at a later stage?
You are free to withdraw from the study at any time, and with no explanation.

What are the risks of taking part in the study?
We do not believe there are any risks in being involved in this study. We appreciate that there may be some inconvenience to you by having to come into the hospital for the discussion but we will try and reduce this by arranging a convenient date and time for you to come.

What are the benefits of taking part in the study?
The benefits of taking part in this study are that you will be helping us to gain a better understanding of stiffness in people with rheumatoid arthritis. This will help us to improve decisions made about treatment and management of rheumatoid arthritis.

Will my taking part in the study be kept confidential?
Yes. Your name will be replaced by a code. All other identifying information (such as people's names, locations or specific descriptions) will be replaced with code numbers or a generalised summary. No one will be able to identify you from any analysis or report. The study reports will include quotations from the discussion but no names will be used. The recordings will be kept securely for 6 years and then destroyed, in accordance with best practice in research guidelines.

What will happen to the results of this research study?
Research team members will analyse the anonymous transcripts and discuss our findings. The findings of this study will influence the design of later research studies within this PhD. We hope the results will be reported in professional journals and at meetings (but participants will not be identified by name). We will send you a summary of the results if you would like.

Who is organising, funding and reviewing the research?
The study is coordinated by a team from the university of the West of England (UWE) based at the Academic Rheumatology Unit at the Bristol Royal Infirmary. It is funded by UWE and has been peer reviewed by the local and UWE Research Ethics Committees.

What do I do now?
Thank you for considering taking part in this research. Please return the reply slip provided if you would like to take part by returning it in the pre-paid reply envelope to Serena Halls. Serena will then contact you in a few days with further information and to answer any questions.

Research team:

Serena Halls, PhD Student Researcher, UWE Bristol
Professor Sarah Hewlett, Professor of Rheumatology Nursing, UWE Bristol
Professor John Kirwan, Professor of Rheumatic Diseases, UoB Bristol
Dr Jon Pollock, Reader in Epidemiology, UWE Bristol
Dr Emma Dures, Research Fellow, UWE Bristol
Mrs Avis Edmunds, Patient Research Partner
Mrs Gill Baker, Patient Research Partner

Contact:

Serena Halls	0117 342 4972	Serena.Halls@uwe.ac.uk
Sarah Hewlett	0117 324 2903	Sarah.Hewlett@uwe.ac.uk

Appendix H: Original focus group topic guide (Study 2)

1. **Check tape recorder works**
2. **Obtain written consent**
3. **Explain the purpose of the session**
4. **Explain the background of stiffness**
5. **Explain ground rules: confidentiality, anonymity, and respect**
6. **Switch tape recorder on!**

Introduction (researchers):

- Researchers introduce themselves
- Thank patients for coming
- Explain purpose of study

Introduction (patients):

- Names/duration of RA

Main body:

***Part A:* The experience of stiffness and how you describe it**

- What is your experience of stiffness? (take notes)
- What words do you use to describe stiffness?
- Liked and disliked descriptors (separate piles)
- What is your definition of stiffness?

***Part B:* Measure development - designing RA stiffness PROM**

- What do you think about these items? (see flipchart)
- What do you like/dislike about them?
- How could they be improved?
- Discuss
    - Wording and format
    - Number of items
    - Timeframe

Debrief:

- Overview and last thoughts
- Thank patients for coming

Other general prompts

Can you give me an example?
Can you explain that a bit more?
Why do you say that?

Appendix I: Revised focus group topic guide (Study 2)

1. **Check tape recorder works**
2. **Obtain written consent**
3. **Explain the purpose of the session**
4. **Explain the background of stiffness**
5. **Explain ground rules: confidentiality, anonymity, and respect**
6. **Switch tape recorder on!**

Introduction (researchers):

- Researchers introduce themselves
- Thank patients for coming
- Explain purpose of study

Introduction (patients):

- Names/duration of RA

Main body:

Part A:* *The experience of stiffness and how you describe it

- What is your experience of stiffness? (take notes)
- What words do you use to describe stiffness? (here or at end)
- Liked and disliked descriptors

Include key discussion point prompts where relevant:

Is stiffness specific?

- Stiffness a patient word
- Stiffness and pain
- Changes in disease activity

Stiffness location

Changes over disease duration

- Change in impact or change in severity?

Part B:* *Measure development - designing RA stiffness PROM

- Study 1 results overview
- Present impact triad concept
- Discuss thoughts on measurement based on aspects of impact triad
- Consider
    - Stem questions
    - Response options
    - Timeframe
    - Layout/format

Debrief:

- Overview and last thoughts
- Thank patients for coming

Appendix J: Study 1 coding tree with Study 2 codes added

Codes	Level 1 codes	Level 2 codes	Theme
Stiffness is normal	Stiffness is normal	Stiffness is a normal consequence of RA	Part of my disease
Stiffness is a normal part of getting older [FG]			
No stiffness prior to RA	No stiffness prior to RA		
Stiffness is an obvious symptom of RA **[reinforced in FG]**	Stiffness is an obvious symptom of RA		
Stiffness is part of a bigger picture in the morning specifically			
Stiffness is part of a bigger picture			
RA stiffness is different – General	RA stiffness is specific		
RA stiffness is different to exercise stiffness			
RA stiffness is different to overuse stiffness			
RA stiffness is different to OA stiffness			
RA stiffness is not different to OA stiffness **[reinforced in FG]**			
Pain is different in RA and OA – but stiffness is the same [FG]			
During a period of flare you cannot move	Stiffness relates to flare	Stiffness varies with disease fluctuations	
During a period of flare it is difficult to move and you can't do ADLs			
Can't do anything in a flare up			
Stiffness affects sleep when it's bad			
During a period of flare pain and stiffness more related			
During a period of flare stiffness and inflammation more related			
During a period of flare all symptoms are worse			
During a period of flare stiffness does not go away			
During a period of flare stiffness sticks around			
During a period of flare stiffness is quick/sudden			
In flare stiffness lasts longer			
In flare stiffness is more frequent in occurrence			
During a period of flare stiffness affects more joints			
During a period of flare stiffness severity is high			

Codes	Level 1 codes	Level 2 codes	Theme
During a period of flare stiffness is an exaggeration of itself			
During a period of flare stiffness is acute			
During a period of flare stiffness you can't use manage it in the same way			
Harder to ease in flare			
Damaged joints cause stiffness	Stiffness relates to damaged joints		
You can't move damaged joints			
Causes of damaged joints			
Stiffness from damage and flare are different			
Pain and stiffness are normal	Patient symptoms – Pain	Relationship between stiffness and other RA symptoms	
Pain and stiffness are related **[reinforced in FG]**			
Pain and stiffness are different concepts **[reinforced in FG]**			
Pain comes with baggage			
Pain is easier to deal with than stiffness			
Pain is harder to deal with than stiffness			
Pain and stiffness more related during flare			
Pain effects sleep			
You feel stiff when you try to move but pain you feel when you don't move			
If you stay in one position for too long you get pain			
You can tell pain and stiffness limitations apart			
Management targets pain and stiffness differently [FG]			
You should separate pain and stiffness [FG]			
You might stiffness as a protective instinct against pain [FG]			
The feeling/sensation of stiffness is a type of pain [FG]			
Hard to define in words the difference between symptoms [FG]			
Uncertainty about the difference between pain and stiffness [FG]			
Stiffness is more physically restrictive than pain [FG]			

Codes	Level 1 codes	Level 2 codes	Theme
Stiffness and fatigue are related	Patient symptoms - Fatigue		
Stiffness is more fatigue than pain			
Stiffness and fatigue are different			
Stiffness and fatigue are distinguishable by the time they occur			
Stiffness is related to inflammation	Medical symptoms – Inflammation		
Stiffness is unrelated to inflammation			
Inflammation causes stiffness			
Inflammation relates to a flare			
Stiffness is the same as inflammation			
Both stiffness and inflammation cause functional loss			
Patients don't recognise inflammation			
All symptoms are intertwined	All symptoms are intertwined		
All symptoms are intertwined when its bad but unrelated when it's not			
All symptoms are worse [in a flare]			
Symptoms are not just related when disease is bad			
Stiffness was an early symptom of RA	Stiffness relevant in early RA	Varying prominence of stiffness during course of the disease	
Stiffness was severe at disease onset			
Stiffness changes over disease duration [FG]			
Stiffness was not relevant in early disease	Stiffness not relevant in early RA		
Stiffness more relevant later in disease duration			
Stiffness affects certain parts of the body	Stiffness affects certain locations	Location within body	Local and widespread
Stiffness feels the same in different part of the body			
Stiffness moves around the body	Stiffness affects variable locations		
Stiffness feels different in different parts of the body			
During a flare/bad day stiffness affects more joints **[reinforced in FG]**	Stiffness all over		
In the morning stiffness affects more joints/all over **[reinforced in FG]**			
Stiffness affects whole body [FG]			
Joints are the problem			

Codes	Level 1 codes	Level 2 codes	Theme
Muscles are the problem	Stiffness affects particular structures	Affected body structures	
Tendons are the problem			
Uncertainty about what structure is the problem	Uncertainty about affected structures		
Stiffness is caused by being immobile	Immobility	Movement and stiffness	Linked to behaviour and environment
Being in a restricted position causes stiffness			
Should be less stiff when have been more active in the night but not the case			
Legacy of activity causes stiffness	Overdoing it		
Overdoing it and not resting causes stiffness			
Medications are beneficial for stiffness – Non-specific	Medications have an impact on stiffness	Medications and stiffness	
Medications are beneficial for stiffness – Infusion			
Medications are beneficial for stiffness – Steroids			
Medications are beneficial for stiffness – DMARD			
Medications are beneficial for stiffness - Anti-TNF			
Medication improvements - now prevent damage **[reinforced in FG]**			
Medications reduce stiffness duration	Medications reduce stiffness		
Medications reduce stiffness severity			
Medications reduce stiffness which allows normality			
Medications get rid of stiffness			
Only steroids reduce stiffness in flare			
Medications reduce stiffness but it's a lost entity because it's never measured	Medications impact on stiffness is lost		
Medications do not target stiffness	Medications are not beneficial for stiffness		
Medications do not work when you have a flare up			
Some medications work better than others	Medications effects can be variable		
Sometimes the same medications work better than other times			
You have transition periods between medications			

Codes	Level 1 codes	Level 2 codes	Theme
Medications have side effects	Medications have other considerations		
Not wanting to take medications			
You don't have a choice but to take medications			
Certain drinks cause stiffness to be worse	Diet	Lifestyle and environment and stiffness	
Certain foods cause stiffness to be worse			
Air conditioning causes joints to be stiff and painful	Weather		
The weather has an impact on stiffness – Hot weather			
The weather has an impact on stiffness – Cold weather			
The weather has an impact on stiffness – Humid weather			
Air pressure has an impact on stiffness [FG]			
The weather does not influence stiffness for everyone [FG]			
Stiffness does not affect sleep	Sleep		
Stiffness does affect sleep **[reinforced in FG]**			
Poor sleep impacts on stiffness			
Stiffness is individual/subjective	Individual experience	Stiffness is individual	Highly variable
Stiffness is different for different people	Different experience		
Stiffness does relate to mornings	Morning	Temporal pattern of stiffness	
Stiffness does not relate to mornings			
You notice stiffness when you get up/wake up			
Stiffness relates to other times of day	Other times of day		
Stiffness doesn't relate to a particular time of day			
Best time of day is afternoon/evening			
Stiffness relates to the night time [FG]			
Timing of stiffness has changed through the course of the disease	Timing changes		
Duration of stiffness – Varies within people	Varies	Duration of stiffness	
Duration of stiffness – Varies between people			
Duration of stiffness - Different amount of time			
Duration of stiffness - Gradually eases off			

Codes	Level 1 codes	Level 2 codes	Theme
Duration of stiffness - Similar amount of time	Constant		
Stiffness never completely goes away			
Stiffness is constant now (it doesn't change)			
Some joints stick all day			
Time is not a relevant factor	Time is not relevant		
General stiffness – Expected	General stiffness	Severity of stiffness	
General stiffness – Impact on function			
General stiffness – Manageable			
Different levels of stiffness **[reinforced in FG]**			
Severe stiffness – Cramp stiffness	Severe stiffness		
Severe stiffness – Not expected [harder to manage]			
Can't function when its bad			
Damage related stiffness – Element of severity in damaged joints	Damage related stiffness		
Can't do ADL's or simple tasks	Impact on activities of daily living and essential tasks	Daily life impact	Impacts on daily life
Restricts hobbies			
Can't eat			
Restricts tasks			
Driving			
Dressing **or undressing [added in FG]**			
Causes difficulties			
Getting up from a chair is difficult			
Stairs			
Restricts family life			
Can't plan			
Difficult to get out of bed			
Unnatural/awkward walking			
Can't get comfortable **[reinforced in FG]**			
Stiffness a problem at night			

Codes	Level 1 codes	Level 2 codes	Theme
Stiffness not a problem at night			
Impact on life			
Restriction			
Disabling			
Stiffness makes everything an effort [FG]			
Affects work	Impact on work		
Loose normality	Impact on normality		
Stiffness in different places is worse because it has different impact	Impact differs in different locations		
Location of stiffness affects the impact [worse in hands]			
Certain movements			
Difficult to move	General impact	Physical impact	
Unable to move			
No dexterity	Specific impact		
Reduced ROM			
Grip			
Bending down			
No balance			
Don’t have mobility			
Slow/can’t rush/actions take me longer			
Physically have to unbend limbs			
No flexibility			
Need physical support			
Loss of strength [FG]			
Forget how to walk properly	Cognitive impact	Cognitive impact	
Have to think about doing actions			
Bits of my body won’t move when I am expecting them to [FG]			
Automatic instinct is gone [FG]			
Psych impact/general wellbeing	Psychological impact	Psychological impact	
Mood affects symptoms			
Stiffness doesn’t affect mood			

Codes	Level 1 codes	Level 2 codes	Theme
Stiffness makes you frustrated**/angry [added in FG]**			
Stress and anxiety impacts on symptoms			
Restricts image – vanity			
You lose the good part of the day			
I worry about stiffness [FG]			
I don't worry about stiffness [FG]			
Stiffness makes me embarrassed [FG]			
Pain on movement	Pain impact	Pain impact	
If you force movements you get pain			
Pain on movement is accentuated in flare			
Must move to relieve stiffness which means you go through the pain barrier			
Stiffness is a warning – telling you to slow down **[reinforced in FG]**			
Moving	General moving	Movement based strategies	Requires self-management
Moving position			
Moving while still in bed			
Uncertainty as to whether movement always helps reduce stiffness [FG]			
Walking	Specific moving		
Stretching			
Supporting joints			
Physically manipulating your joints			
Balancing rest and movement			
Exercise	Exercise		
Debate about the benefit of exercise [FG]			
Gadgets/aids	Gadgets and splints	Other strategies	
Splints are effective for stiffness			
Splints are not effective for stiffness			
Splints are not compatible with other self-management strategies			
Hot/cold therapy is effective	Heat and cold		

Codes	Level 1 codes	Level 2 codes	Theme
Shower in the morning			
Hot/cold therapy is not effective			
Hydrotherapy			
Alternative therapy - Changing diet	Alternative therapies		
Alternative therapy - Aromatherapy			
Alternative therapy - Relaxation techniques			
Take medication in the morning to get going	Medications and painkillers		
Take medication to function			
Take painkillers			
Family and friends – Physically help with jobs/housework	External – Social support	Psychosocial strategies	
Family and friends - Facilitate use of self-management strategies			
Family and friends - Are flexible			
Normalise/accept stiffness	Internal – Normalise		
Adapt and adjust behaviours generally - Have to find other ways to do things	Internal – Adapt and adjust		
Adapt and adjust behaviours generally - You can work round stiffness			
Adapt and adjust behaviours in the morning - You just adapt in the morning			
Adapt and adjust behaviours in the morning - Perform activities later in day			
Adapt and adjust behaviours in the morning - Use gadgets at difficult times			
Prepare/plan for stiffness - Compensate for slow movement by getting up earlier	Internal – Prepare and plan		
Prepare/plan for stiffness - Not too restrictive if pace and plan			
Self-management is easier since stopping work			
Ensure RA and stiffness does not take over life [RA general]	Internal – Part of general RA management		
Self-management is individual/knowing your limitations/your body [general RA]			
Do what you can to help yourself [general RA]			

Codes	Level 1 codes	Level 2 codes	Theme
Self-management and understanding RA develops over time [RA general]			
Self-management can be difficult especially when you enjoy things			

Appendix K: Study 2 coding tree

Coding frame	Level 1 codes	Level 2 codes	Level 3 codes	Key areas
Stem questions and anchors	Relevant to the individual	Stiffness is very individual and measurement should reflect that		Stem questions and anchors: Individual
		Questions should be worded around the individual (“to you”) to appreciate whether this is usual		
		Questions need to consider individual response shift		Stem questions and anchors: Individual; Response shift
		Wording questions around the individual can be difficult because of variability		Stem questions and anchors: Individual
	Impact	Relevant consequences are just normal everyday tasks		Stem questions and anchors: Impact
		Possible impact questions worded around daily tasks or movement		
		Impact around taking longer to do things		
		Comes down to impact on QoL		
		Difficulties with impact questions	Some impact questions are gender specific	
			Capture difficulty not completion	
			Time and location specific	
		Liked the impact triad as a concept around which to base measurement	Impact triad mirrors real life experiences	Stem questions and anchors: Impact triad
			Impact triad accounts for individual experience	

Coding frame	Level 1 codes	Level 2 codes	Level 3 codes	Key areas
	Stiffness after a period of immobility			Stem questions and anchors: Stiffness after period of immobility
	Timing and temporal pattern	Time is relevant		Stem questions and anchors: Duration is relevant but difficult
		Current duration questions are difficult answer	Difficult to quantify	
			It is hard to remember	
			What is the start and endpoint?	
			It depends what you are doing	
			Duration of other symptoms not considered	
		Questions about morning stiffness are limited	Morning stiffness is not relevant to patients	Stem questions and anchors: When do you get stiff is better to ask
			Morning stiffness is important but should be separated from daytime stiffness	
			Difficulties answering questions about the morning	
		When is a more appropriate question	Need to ask when stiffness occurs rather than about the morning?	
			Need to ask about stiffness during the night	
	Location	Ask where you are stiff	Number of joints affected	Stem questions and anchors: Where do you get stiff
			Reflects individual experience	
			Relates to usual experience	
			You could look at structure and location this way	

Coding frame	Level 1 codes	Level 2 codes	Level 3 codes	Key areas
		Location relates to severity	Could ask both where and how bad	
			Assessment of more joints does not indicate and increase in severity for everyone	
		Assess each joint individually for severity?	Different joints might have different severities	
	Pain, stiffness and other symptoms	Would like to be asked about pain and stiffness but it must be separate		Stem questions and anchors: Other symptoms (esp. pain)
		Need to ask about pain and stiffness separately or it is confusing		
		Asking about 'painful stiffness' separates mechanical stiffness from inflammatory stiffness		
		Ask about pain, inflammation and stiffness		
Response options	Difference in opinion about response options	VAS lines are imprecise		Response options: Difference in opinion about VAS/NRS/VRS
		VAS are easy to respond to		
		NRS - numbers are preferable to lines		
		NRS provide too many options		
		Less options are better		
	Response option preference	Fewer response options		
		Free text		
Timeframe	Debate about appropriate timeframe	Timeframe must be recent		Timeframe: Debate about appropriate timeframe
		Stiffness over the last week con - might not capture the worst times/daily variability		
		Stiffness over the last week pro – acceptable and accurate		

Coding frame	Level 1 codes	Level 2 codes	Level 3 codes	Key areas
		Stiffness now pro – takes a snapshot of you now		
		Stiffness now pro – captures the worst for those on direct access		
		Stiffness now con – might not capture the worst times/variability		
Layout and format	Patients want to know why you are asking	Provide explanation to improve accuracy		Layout and format: Purpose
	Visual aspects might be effective	Clock face		Layout and format: Visual elements
		Body image		
	It must be short and simple	Give examples to help simplify		Layout and format: Succinct
		Fewer response options		
	Questions must be up to date and not old fashioned			Layout and format: Modern
	Debate about method of input/format	Room for free text boxes to make it individual		Layout and format: Practicality of different approaches
		Practicality of lists – you need to use general categories to shorten the process		
		Words to circle or tick are better than lots to write		
		Circles are hard to draw		
	Specific wording	Joint stiffness?		Layout and format: Specific wording
		Flare vs seize up stiffness		
		Severity or level		
		Stiffness in every stem?		

Appendix L: COREQ checklist (Study 2)

No. Item	Guide questions/description	How was this component addressed
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Interviewer	Which author/s conducted the interview or focus group?	Identified in Section 5.3.4
2. Credentials	What were the researcher's credentials?	The characteristics of the supervisory team were identified in Section 3.4.2 and the researchers (Halls) background was highlighted in Section 3.2.1
3. Occupation	What was their occupation at the time of the study?	
4. Gender	Was the researcher male or female?	
5. Experience and training	What experience or training did the researcher have?	
Relationship with participants		
6. Relationship established	Was a relationship established prior to study commencement?	
7. Participant knowledge of the interviewer	What did the participants know about the researcher?	Identified in Section 5.3.4
8. Interviewer characteristics	What characteristics were reported about the interviewer?	The researchers (Halls) background, experience and research interests were highlighted in Section 3.2.1
Domain 2: study design		
Theoretical framework		
9. Methodological orientation and theory	What methodological orientation was stated to underpin the study?	The methodological approach to the study was discussed in Section 3.3.1 and the method and analysis approach were discussed in Sections 5.3.1 and 5.4
Participant selection		
10. Sampling	How were participants selected?	
11. Method of approach	How were participants approached?	Identified in Section 5.3.2
12. Sample size	How many participants were in the study?	
13. Non-participation	How many people refused to participate or dropped out? Reasons?	Identified in Section 5.5.1
Setting		
14. Setting of data collection	Where was the data collected?	Identified in Section 5.3.4

15. Presence of non-participants	Was anyone else present besides the participants and researchers?	NA
16. Description of sample	What are the important characteristics of the sample?	Identified in Section 5.5.1
Data collection		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Identified in Section 5.3.3 and Appendices H and I
18. Repeat interviews	Were repeat interviews carried out? If yes, how many?	This was not performed
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	Identified in Section 5.3.4
20. Field notes	Were field notes made during and/or after the interview or focus group?	Identified in Section 5.4
21. Duration	What was the duration of the interviews or focus group?	Identified in Section 5.5.2
22. Data saturation	Was data saturation discussed?	Identified in Section 5.3.2
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	This was not performed
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	Identified in Section 5.4
25. Description of the coding tree	Did authors provide a description of the coding tree?	Identified in Sections 5.5.2.2, 5.5.2.4, and Appendices J and K
26. Derivation of themes	Were themes identified in advance or derived from the data?	Identified in Section 5.4
27. Software	What software, if applicable, was used to manage the data?	Identified in Section 5.4
28. Participant checking	Did participants provide feedback on the findings?	This was not performed
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes? Was each quotation identified?	
30. Data and findings consistent	Was there consistency between the data presented and the findings?	Identified in Section 5.5
31. Clarity of major themes	Were major themes clearly presented in the findings?	
32. Clarity of minor themes	Is there a description of diverse cases or minor themes?	

Appendix M: Identifying what is relevant to measure

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
Part of having RA	Stiffness is a normal consequence of RA	Stiffness is normal	Stiffness is normal	As a result of stiffness being considered to be normal in RA and part of having RA coupled with stiffness also being relevant in other conditions such as OA, the use of specific wording may be important. This would also aid keeping the focus on stiffness during completion of the questionnaire	Part of having RA: Normal and specific nature
			Stiffness is a normal part of getting older [FG]		
		No stiffness prior to RA	No stiffness prior to RA		
		Stiffness is an obvious symptom of RA	Stiffness is an obvious symptom of RA [reinforced in FG]		
			Stiffness is part of a bigger picture in the morning specifically		
			Stiffness is part of a bigger picture		
		RA stiffness is specific	RA stiffness is different – General		
			RA stiffness is different to exercise stiffness		
			RA stiffness is different to overuse stiffness		
			RA stiffness is different to OA stiffness		
			RA stiffness is not different to OA stiffness [reinforced in FG]		
			Pain is different in RA and OA – but stiffness is the same [FG]		
	Stiffness varies with disease fluctuations	Stiffness relates to flare	During a period of flare you cannot move	As there was an indication that stiffness was influenced by fluctuating disease it would be important to capture this change in stiffness in measurement	Part of having RA: Fluctuation with disease
			During a period of flare it is difficult to move and you can't do ADLs		
			Can't do anything in a flare up		
			Stiffness affects sleep when it's bad		
			During a period of flare pain and stiffness more related		
			During a period of flare stiffness and inflammation more related		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
			During a period of flare all symptoms are worse		
			During a period of flare stiffness does not go away		
			During a period of flare stiffness sticks around		
			During a period of flare stiffness is quick/sudden		
			In flare stiffness lasts longer		
			In flare stiffness is more frequent in occurrence		
			During a period of flare stiffness affects more joints		
			During a period of flare stiffness severity is high		
			During a period of flare stiffness is an exaggeration of itself		
			During a period of flare stiffness is acute		
			During a period of flare stiffness you can't use manage it in the same way		
			Harder to ease in flare		
		Stiffness relates to damaged joints	Damaged joints cause stiffness	It would also be important to differentiate stiffness as a result of different processes	Part of having RA: Process
			You can't move damaged joints		
			Causes of damaged joints		
			Stiffness from damage and flare are different		
	Relationship between stiffness and other RA symptoms	Patient symptoms – Pain	Pain and stiffness are normal	The relationship between stiffness and other symptoms was clear and this must be considered in instrument	Part of having RA: Relationship with other symptoms
			Pain and stiffness are related [reinforced in FG]		
			Pain and stiffness are different concepts [reinforced in FG]		
			Pain comes with baggage		
			Pain is easier to deal with than stiffness		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
			Pain is harder to deal with than stiffness	development. As above the use of specific language will ensure that the topic of stiffness is reinforced. Other symptoms will be assessed as part of the questionnaire pack for validity testing	
			Pain and stiffness more related during flare		
			Pain effects sleep		
			You feel stiff when you try to move but pain you feel when you don't move		
			If you stay in one position for too long you get pain		
			You can tell pain and stiffness limitations apart		
			Management targets pain and stiffness differently [FG]		
			You should separate pain and stiffness [FG]		
			You might stiffness as a protective instinct against pain [FG]		
			The feeling/sensation of stiffness is a type of pain [FG]		
			Hard to define in words the difference between symptoms [FG]		
			Uncertainty about the difference between pain and stiffness [FG]		
			Stiffness is more physically restrictive than pain [FG]		
		Patient symptoms - Fatigue	Stiffness and fatigue are related		
			Stiffness is more fatigue than pain		
			Stiffness and fatigue are different		
			Stiffness and fatigue are distinguishable by the time they occur		
		Medical symptoms – Inflammation	Stiffness is related to inflammation		
			Stiffness is unrelated to inflammation		
			Inflammation causes stiffness		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
			Inflammation relates to a flare		
			Stiffness is the same as inflammation		
			Both stiffness and inflammation cause functional loss		
			Patients don't recognise inflammation		
		All symptoms are intertwined	All symptoms are intertwined		
			All symptoms are intertwined when its bad but unrelated when it's not		
			All symptoms are worse [in a flare]		
			Symptoms are not just related when disease is bad		
	Varying prominence of stiffness during course of the disease	Stiffness relevant in early RA	Stiffness was an early symptom of RA	Could explore differences in response between newly diagnosed vs experienced patients in Study 4	
			Stiffness was severe at disease onset		
			Stiffness changes over disease duration [FG]		
		Stiffness not relevant in early RA	Stiffness was not relevant in early disease		
			Stiffness more relevant later in disease duration		
Local and widespread	Location within body	Stiffness affects certain locations	Stiffness affects certain parts of the body	This may capture aspects of severity and/or impact that have not been tested before	Local and widespread: Location
			Stiffness feels the same in different part of the body		
		Stiffness affects variable locations	Stiffness moves around the body		
			Stiffness feels different in different parts of the body		
		Stiffness all over	During a flare/bad day stiffness affects more joints [reinforced in FG]		
			In the morning stiffness affects more joints/all over [reinforced in FG]		
			Stiffness affects whole body [FG]		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
	Affected body structures	Stiffness affects particular structures	Joints are the problem	This may be relevant from a severity perspective and also in stiffness assessment across conditions e.g. PMR – stiff in muscles	Local and widespread: Structure
			Muscles are the problem		
			Tendons are the problem		
		Uncertainty about affected structures	Uncertainty about what structure is the problem		
Linked to behaviour and environment	Movement and stiffness	Immobility	Stiffness is caused by being immobile	Important to consider stiffness after immobility, and was patient relevant	Linked to behaviour and environment: Movement
			Being in a restricted position causes stiffness		
			Should be less stiff when have been more active in the night but not the case		
		Overdoing it	Legacy of activity causes stiffness		
			Overdoing it and not resting causes stiffness		
	Medications and stiffness	Medications have an impact on stiffness	Medications are beneficial for stiffness – Non-specific	Patient relevant to discuss medications. If this is the way that some patients assess stiffness then it may be important to test. Could consider about medications are working as effectively as usual? Although this may	Linked to behaviour and environment: Medication effectiveness
			Medications are beneficial for stiffness – Infusion		
			Medications are beneficial for stiffness – Steroids		
			Medications are beneficial for stiffness – DMARD		
			Medications are beneficial for stiffness - Anti-TNF		
			Medications are better than they were - now prevent damage [reinforced in FG]		
		Medications reduce stiffness	Medications reduce stiffness duration		
			Medications reduce stiffness severity		
			Medications reduce stiffness which allows normality		
			Medications get rid of stiffness		
			Only steroids reduce stiffness in flare		
		Medications impact on	Medications reduce stiffness but it's a lost entity because it's never measured		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
		stiffness is lost		not need direct assessment	
		Medications are not beneficial for stiffness	Medications do not target stiffness		
			Medications do not work when you have a flare up		
		Medications effects can be variable	Some medications work better than others		
			Sometimes the same medications work better than other times		
			You have transition periods between medications		
		Medications have other considerations	Medications have side effects	Purely experiential and not relevant to measurement	
			Not wanting to take medications		
			You don't have a choice but to take medications		
	Lifestyle and environment and stiffness	Diet	Certain drinks cause stiffness to be worse	As with medications, if these are the ways that patients use to talk about and assess stiffness then they may need consideration in assessment. Although again this may not need direct assessment	Linked to behaviour and environment: Effect of lifestyle and environment
			Certain foods cause stiffness to be worse		
		Weather	Air conditioning causes joints to be stiff and painful		
			The weather has an impact on stiffness – Hot weather		
			The weather has an impact on stiffness – Cold weather		
			The weather has an impact on stiffness – Humid weather		
			Air pressure has an impact on stiffness [FG]		
			The weather does not influence stiffness for everyone [FG]		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
		Sleep	Stiffness does not affect sleep		
			Stiffness does affect sleep [reinforced in FG]		
			Poor sleep impacts on stiffness		
Highly variable	Stiffness is individual	Individual experience	Stiffness is individual/subjective	This was very important to patients and reinforced by patient partners. This may influence how questions are worded	Highly variable: Individual experience
		Different experience	Stiffness is different for different people		
	Temporal pattern of stiffness	Morning	Stiffness does relate to mornings	Key to capture temporality in assessment bearing in mind patient experience being broader than the morning and duration being difficult	Highly variable: Temporal pattern
			Stiffness does not relate to mornings		
			You notice stiffness when you get up/wake up		
		Other times of day	Stiffness relates to other times of day		
			Stiffness doesn't relate to a particular time of day		
			Best time of day is afternoon/evening		
			Stiffness relates to the night time [FG]		
		Timing changes	Timing of stiffness has changed through the course of the disease		
	Duration of stiffness	Varies	Duration of stiffness – Varies within people	See above. Plus could consider assessing duration in relation to impact e.g. it takes me long to complete task/s	Highly variable: Duration
			Duration of stiffness – Varies between people		
			Duration of stiffness - Different amount of time		
			Duration of stiffness - Gradually eases off		
		Constant	Duration of stiffness - Similar amount of time		
			Stiffness never completely goes away		
			Stiffness is constant now (it doesn't change)		
			Some joints stick all day		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
		Time is not relevant	Time is not a relevant factor		
	Severity of stiffness	General stiffness	General stiffness – Expected	Important to capture severity in assessment considering the usual baseline level and maybe severity in relation to impact	Highly variable: Severity
			General stiffness – Impact on function		
			General stiffness – Manageable		
			Different levels of stiffness [reinforced in FG]		
		Severe stiffness	Severe stiffness – Cramp stiffness		
			Severe stiffness – Not expected [harder to manage]		
			Can't function when its bad		
		Damage related stiffness	Damage related stiffness – Element of severity in damaged joints		
Impacts on daily life	Daily life impact	Impact on activities of daily living and essential tasks	Can't do ADL's or simple tasks	The key aspect from the interviews and focus groups and a very natural way for patients to discuss stiffness. Important to ensure that every aspect here is captured and tested	Impacts on daily life: Daily life impact
			Restricts hobbies		
			Can't eat		
			Restricts tasks		
			Driving		
			Dressing or undressing [added in FG]		
			Causes difficulties		
			Getting up from a chair is difficult		
			Stairs		
			Restricts family life		
			Can't plan		
			Difficult to get out of bed		
			Unnatural/awkward walking		
			Can't get comfortable [reinforced in FG]		
			Stiffness a problem at night		
			Stiffness not a problem at night		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
			Impact on life		
			Restriction		
			Disabling		
			Stiffness makes everything an effort [FG]		
		Impact on work	Affects work		
		Impact on normality	Loose normality		
		Impact differs in different locations	Stiffness in different places is worse because it has different impact		
			Location of stiffness affects the impact [worse in hands]		
			Certain movements		
	Physical impact	General impact	Difficult to move	As above	Impacts on daily life: Physical impact
			Unable to move		
		Specific impact	No dexterity		
			Reduced ROM		
			Grip		
			Bending down		
			No balance		
			Don’t have mobility		
			Slow/can’t rush/actions take me longer		
			Physically have to unbend limbs		
			No flexibility		
			Need physical support		
			Loss of strength [FG]		
	Cognitive impact	Cognitive impact	Forget how to walk properly	As above	Impacts on
			Have to think about doing actions		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
			Bits of my body won't move when I am expecting them to [FG]		daily life: Cognitive impact
			Automatic instinct is gone [FG]		
	Psychological impact	Psychological impact	Psych impact/general wellbeing	As above	Impacts on daily life: Psychosocial impact
			Mood affects symptoms		
			Stiffness doesn't affect mood		
			Stiffness makes you frustrated/angry [added in FG]		
			Stress and anxiety impacts on symptoms		
			Restricts image – vanity		
			You lose the good part of the day		
			I worry about stiffness [FG]		
			I don't worry about stiffness [FG]		
			Stiffness makes me embarrassed [FG]		
	Pain impact	Pain impact	Pain on movement	This may be an idea to consider however given the relationship between stiffness and pain this may be difficult	Impacts on daily life: Pain impact
			If you force movements you get pain		
			Pain on movement is accentuated in flare		
			Must move to relieve stiffness which means you go through the pain barrier		
			Stiffness is a warning – telling you to slow down [reinforced in FG]		
Requires self-management	Movement based strategies	General moving	Moving	Important to capture this in assessment as was relevant to patients. It also forms part of the impact triad so can be tested within that concept also	Requires self-management: Direct
			Moving position		
			Moving while still in bed		
			Uncertainty as to whether movement always helps reduce stiffness [FG]		
		Specific moving	Walking		
			Stretching		
			Supporting joints		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
			Physically manipulating your joints		
			Balancing rest and movement		
		Exercise	Exercise		
			Debate about the benefit of exercise [FG]		
	Other strategies	Gadgets and splints	Gadgets/aids		
			Splints are effective for stiffness		
			Splints are not effective for stiffness		
			Splints are not compatible with other self-management strategies		
		Heat and cold	Hot/cold therapy is effective		
			Shower in the morning		
			Hot/cold therapy is not effective		
			Hydrotherapy		
		Alternative therapies	Alternative therapy - Changing diet		
			Alternative therapy - Aromatherapy		
			Alternative therapy - Relaxation techniques		
		Medications and painkillers	Take medication in the morning to get going		
			Take medication to function		
			Take painkillers		
	Psychosocial strategies	External – Social support	Family and friends – Physically help with jobs/housework	As above	Requires self-management: Indirect
			Family and friends - Facilitate use of self-management strategies		
			Family and friends - Are flexible		
		Internal – Normalise	Normalise/accept stiffness		
			Adapt and adjust behaviours generally - Have to find other ways to do things		

Theme	Level 2 groups	Level 1 groups	Codes	Relevant to measure discussion	Key areas
		Internal – Adapt and adjust	Adapt and adjust behaviours generally - You can work round stiffness		
			Adapt and adjust behaviours in the morning - You just adapt in the morning		
			Adapt and adjust behaviours in the morning - Perform activities later in day		
			Adapt and adjust behaviours in the morning - Use gadgets at difficult times		
		Internal – Prepare and plan	Prepare/plan for stiffness - Compensate for slow movement by getting up earlier		
			Prepare/plan for stiffness - Not too restrictive if pace and plan		
			Self-management is easier since stopping work		
		Internal – Part of general RA management	Ensure RA and stiffness does not take over life [RA general]	Purely experiential and not relevant to measurement	
			Self-management is individual/knowing your limitations/your body [general RA]		
			Do what you can to help yourself [general RA]		
			Self-management and understanding RA develops over time [RA general]		
			Self-management can be difficult especially when you enjoy things		

Appendix N: Mapping the experience of stiffness to the measurement perspective

Key areas in the experience of stiffness	Key areas regarding stiffness assessment	Summary of measurement ideas	Discussion with team
Part of having RA: Normal and specific nature	Layout and format: Specific wording	A suggestion was made about specifying stiffness in each question to reinforce the topic of the question/s	• Given the complexity of stiffness and how it fits within a complex disease it is important to include stiffness in every stem to avoid distraction and keep on topic? • Keep stiffness or RA stiffness in every stem [JK]
Part of having RA: Fluctuation with disease			• Capture via specific aspects of stiffness in periods of fluctuating disease (detailed in Appendix L)
Part of having RA: Process	Layout and format: Specific wording	Suggested differences between stiffness from different processes. One suggestion about different wording for different experiences of stiffness	• Any different wording to distinguish mechanical/inflammatory? [JK] e.g. constant (fused joints) [ED] • Consider what is used clinically e.g. physio - distinction between active and passive movements -would this or something similar work? [JK] • Permanence-this never moves vs. recent/new/past 7 days? [SHa] Fits with AE experience • Include a distinction for patients about what this is including and excluding e.g. stiffness that comes and goes not damaged joints (clinically HPs do this all the time active vs. passive) [JK]

Key areas in the experience of stiffness	Key areas regarding stiffness assessment	Summary of measurement ideas	Discussion with team
Part of having RA: Relationship with other symptoms	Stem questions and anchors: Other symptoms (esp. pain) Layout and format: Visual elements	Suggestions about pain and stiffness in particular included asking about both but separately to avoid confusion. Plus the idea of 'painful stiffness' was suggested as relevant- could this separate different types of stiffness (Part of having RA: Mechanical vs. inflammatory stiffness)? Also it was suggested that inflammation was also relevant to ask about in addition to pain and stiffness. This lead to the suggestion of what contributes to your experience of stiffness. Visual aspects were encouraged within focus group	• Combined pain and stiffness questions not ideal for measurement and difficult for patients to answer. Therefore need to think about what makes measurement sense here [SH] • Painful stiffness suggested in FG's [ED] • Need to consider the complex relationship between pain and stiffness but bear in mind the purpose of this PROM to assess stiffness • Relevant to measurement because will asking about stiffness capture anything over and above that already captured by pain? [JK] • Related to the idea below (see Location) about capturing painful, swollen and stiff joints on a body and then seeing whether that creates another variable. This would allow you to explore the relationship between all combinations of the above. Although difficulties with patients accurately identifying swollen joints - Sarah's previous work, might influence suggestion about identifying swollen, painful and stiff joints [JK]

Key areas in the experience of stiffness	Key areas regarding stiffness assessment	Summary of measurement ideas	Discussion with team
Local and widespread: Location	Stem questions and anchors: Where do you get stiff Layout and format: Visual elements	This was suggested as a natural way of thinking about stiffness. It was suggested as an indicator of worsening stiffness (if you know where you usually get stiff then if it is broader than usual then that is an indicator of severity (e.g. number of joints affected). However it was highlighted that this was not an indicator of severity for everyone! (See Severity below) Visual aspects were encouraged within focus group	• We currently assess inflammation by number of painful and swollen joints. Number of stiff joints has not been done before but there is precedence [SH] • AE's experience reinforced the idea of overall stiffness when struggling • Interesting idea, suggested trying the visual representation of stiffness with patients [might be different with different patients e.g. RA vs. PMR] [JK] • All over stiffness option - similar to body used in the PDAS [ED] • Another way of doing this would be ask where do you feel stiff-one joint/two joints/outside the joints/all joints? Maybe look for patterns (similar to McGill pain questionnaire). Overall, location of stiffness worth exploring [JK]
Local and widespread: Structure	Stem questions and anchors: Where do you get stiff Layout and format: Specific wording	If you asked where you might get a better idea of the affected structure Suggestion about stiffness everywhere A suggestion was made to specify 'joint stiffness'	• AE's experience reinforced the idea of overall stiffness when struggling • All over stiffness option - similar to body used in the PDAS [ED] • Include specific wording

Key areas in the experience of stiffness	Key areas regarding stiffness assessment	Summary of measurement ideas	Discussion with team
Linked to behaviour and environment: Movement	Stem questions and anchors: Stiffness after period of immobility	This is something that is relevant to patients but currently not being captured in current measurement. Could this separate different types of stiffness	• Capture stiffness after a period of immobility • Could asking about whether stiffness reduces following exercise/movement differentiate mechanical and inflammatory stiffness [JP]
Linked to behaviour and environment: Medication effectiveness			• Patient relevant to capture aspects like whether medications are working effectively/as usual?
Linked to behaviour and environment: Effect of lifestyle and environment			• Patient relevant to capture aspects like weather, diet, sleep?
Highly variable: Individual experience	Stem questions and anchors: Individual Stem questions and anchors: Individual; Response shift	It was highlighted that questions need to appreciate individual differences and reflect whether what you have experienced in the timeframe is usual for you. However this may be difficult given the variability of stiffness. Response shift highlighted as being an influence on perceptions. Maybe this could be explored somehow?	• Look at wording in the PFQ worded around normal/usual/average for me [SH] • Include individual wording in wording questions • Capture response shift?

Key areas in the experience of stiffness	Key areas regarding stiffness assessment	Summary of measurement ideas	Discussion with team
Highly variable: Temporal pattern	Stem questions and anchors: When do you get stiff is better to ask Layout and format: Visual elements	The concept of stiffness is broader than just morning stiffness for most patients (although morning stiffness is relevant, particularly for some). The morning is difficult to answer about (e.g. what is the morning?). Stiffness during the night is also relevant. Maybe WHEN is a better way to ask? Visual aspects were suggested within focus group	• AE suggested that using the term morning stiffness excludes other relevant times and liked the idea about when • Stiffness at night challenges the concept of EMS [ED]. • Timing graph or clock face as visual options - Could measure either using area under the curve, digitise it? Lots of possibilities but also lots of challenges. Overall, some visual way of looking at this sounds possible but need to find out what is feasible for patients (OMERACT filter) [JK] • Related to how far off normal am I? [ED]
Highly variable: Duration	Stem questions and anchors: Duration is relevant but difficult Stem questions and anchors: When do you get stiff is better to ask	It was suggested that duration is relevant to patients but it is difficult to answer; it doesn't capture the whole stiffness experience, we don't ask how long pain lasts, it is hard to remember and put a time on it, we don't sit and time it, it depends what you are doing, what are the start and end points etc. Lots of suggestions about how to combat this Maybe WHEN is a better way to ask?	• As above – when are you stiff? • Address through impact e.g. does it take longer to complete certain tasks?

Key areas in the experience of stiffness	Key areas regarding stiffness assessment	Summary of measurement ideas	Discussion with team
Highly variable: Severity	Stem questions and anchors: Where do you get stiff	it was suggested that number of affected joints may indicate severity (See Location above)	• Address through impact e.g. stiffness is so severe I am unable to complete certain tasks • Capture the number of affected joints (See Location above) • Consider usual baseline?
Impacts on daily life: Daily life impact Impacts on daily life: Physical impact Impacts on daily life: Cognitive impact Impacts on daily life: Psychosocial impact Impacts on daily life: Pain impact	Stem questions and anchors: Impact Stem questions and anchors: Impact triad Layout and format: Specific wording	There were suggestions that impact relates to normal everyday tasks and activities. Specific examples were identified (captured in inductive analysis) and impact was also related to taking longer to do tasks as well as them being more difficult. Difficulties with impact questions were highlighted including certain impact elements being gender, time (e.g. morning), or location (e.g. affected body part) specific, and the input of others (e.g. help from significant other). The impact triad was seen as a relevant concept to patients to base measurement around. Each aspect of the triad was discussed and was relevant. Indicating that the impact triad was acceptable to patients as a concept on which to base the measurement of stiffness. Specific wording was also suggested about how to word different aspects of the triad	• The idea of the impact triad being acceptable to patients is not surprising as it was developed by patients [JK] • Captured as part of the impact triad • Test items capturing each suggested aspect of each area of impact

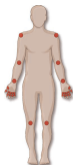
Key areas in the experience of stiffness	Key areas regarding stiffness assessment	Summary of measurement ideas	Discussion with team
Requires self-management: Direct			• Captured as part of the impact triad
Requires self-management: Indirect			• Captured as part of the impact triad
Requires self-management: Change in ability to manage			• Captured as part of the impact triad • Other?
	Response options: Difference in opinion about VAS/NRS/VRS	As there was difference of opinion as to the most appropriate format, different formats could be tested in cognitive interviews and even in the survey. Also relevant to consider aspects highlighted in Format: Practicality of different approaches (see below) e.g. circle are difficult	• Could test questions with different response options?
	Timeframe: Debate about appropriate timeframe	There was debate about the timeframe that was most relevant to stiffness, weighed up against what could be reasonably captured in a questionnaire	• Past 7 days preferred by AE • Could test questions with different timeframes? • Consider the usefulness from a clinical context • Last week is the best for people-captures weekly patterns/easy to remember but given variability check that it is do-able for patients [JK]

Key areas in the experience of stiffness	Key areas regarding stiffness assessment	Summary of measurement ideas	Discussion with team
	Layout and format: Purpose	As it was suggested that there was uncertainty about why this information was being asked for and what it would be used for a brief introduction could help explain the purpose of the question/s	• Include introduction to describe purpose
	Layout and format: Succinct	Must ensure that the questionnaire is simple, as short as possible and that each question only has as many options as necessary. Also provide examples if needed to enhance clarity	• Overall aim is to identify the smallest combination of items that work most effectively together
	Layout and format: Modern	Ensure that questions are relevant and modern (not old-fashioned)	• Consider qualitative data to word items
	Layout and format: Practicality of different approaches	Relevant to the layout of the question/s e.g. free text options, and also to Response options (see above) e.g. circles are hard to draw	

Appendix O: Early item development

Key area	Specific aspect to capture	Potential wording	Origin of wording	Discussion with team
Part of having RA: Normal and specific nature	Specificity	No direct question but include the words stiffness and or RA in every stem question?	QD, SH, JK, GB	Important to retain focus on topic of stiffness [JK/SH/GB]. Could use 'stiffness due to your RA' [JK], 'RA stiffness' [SHa], or "thinking about stiffness" [GB]
	Normality	It was acknowledged that stiffness was part of the 'normal' experience of RA thus maybe we need to capture when stiffness is different to normal?	QD, SH, JK	Taking account of normal adaption was important to patients in development of PFQ (although didn't get into final items) [SH]. This might be captured by addressing the response shift i.e. how patients have reset values. However this would have to separate this from the other item so it doesn't influence responses [SH/JK]. Could ask the level of stiffness on a normal day using 0-10 [SH] and then use it to standardise response [JK]
Part of having RA: Fluctuation with disease	Change	Capture change in impact? E.g. unable to function because of stiffness	SH, JK	No specific question - this will be captured in the impact triad (impact and management) and depending on the response options for impact items this may be captured there. Acknowledge variation through timescale – From a research/clinical perspective we would be looking at the effects of an intervention over weeks/months rather than a few days/hours [JK/SH]. Plus from a recall perspective 7 days is easy to remember as it has a natural social cycle [JK]
		Capture change in self-management?		
		Is change captured in other items e.g. impact items?		
Part of having RA: Process	Process	Do you have some joints that are always/permanently stiff and do not move? (Permanent vs. variable)	JK, ED, QD	Must consider how this information would be used [SH] Consider something like 'I have some joints that do not move normally whether my arthritis is good or bad' [JK]. Always difficult to move [SH] rather than permanent The purpose here is to identify mechanical changes in the joint that are stopping that joint from moving fully to say to patients – we already know about this and we don't need
		I have some joints that do not move even when I try to move them? (Passive vs active)	JK, ED, QD	
		Are you able to reduce stiffness by exercising or moving? (Permanent vs. variable)	QD	
		It is usual for some of my joints not to move even when I try to move them? (Passive vs active: usual)	JK, ED, QD	

Key area	Specific aspect to capture	Potential wording	Origin of wording	Discussion with team
		It is usual for some of my joints not to move permanently? (Permanent vs. variable: usual)	QD, ED, QD	any more information about it - so somehow I think you do have to try and either identify it or get patients in their mind to identify it but then discount it. Maybe have an introduction that introduces these ideas and then a question at the end that checks this e.g. how much of the stiffness you have reported in these questions do you think is due to permanently damaged joints bracket? Having it at the end will mean it doesn't influence the rest of the answers [JK]
		Statement rather than direct question e.g. "We recognise that sometimes joints do get damaged and as a result do not move very much even when you try to move them. In this questionnaire, we are trying to understand your disease activity rather than the consequences of long term disease, so please a) only complete the following questions about your stiffness that comes and goes or changes (i.e. not permanent stiffness) b) do not include stiffness as a result of damage in your answers to the following questions". Could be included with other questions e.g. "If you answered yes to any of the questions above please refer to statement"	QD	
Part of having RA: Relationship with other symptoms	Pain	Ask about painful stiffness e.g. In the past [insert timeframe] have you experienced painful stiffness?	QD, GB, ED	Although this was suggested by patients, Gill suggested avoiding direct discussion about other symptoms to keep the focus on stiffness [GB]. Also from a measurement perspective this is probably sensible to avoid [JP] plus pain will be assessed for validity testing.
	Pain	Capture whether stiffness relates to other symptoms more than usual to pull out increased association between stiffness and other symptoms in flare (inflammatory)	QD	Using 'symptoms' would be good to capture other symptoms rather than just pain – can be interpreted in any way by the patient [JK]. This could be a way of characterising the nature of stiffness [JK]
Part of having RA: Relationship with other symptoms	Inflammation	Do you experience stiffness with inflammation/swelling?	QD	As above, capturing whether stiffness is associated with other 'symptoms' may be useful [JK]. There could be a relevant variable that emerges from the combination of all three? [SH]. But there may be difficulty capturing swollen joints visually or otherwise is that it has been shown that patients can't identify swollen joints accurately (Hewlett, 1995) [SH/JK]
		Do you feel stiff because your joints are swollen?	QD	
		Do your joints feel tight?	QD	
	Combination	Capture stiff, swollen and painful joints visually e.g.	QD, SH	

Key area	Specific aspect to capture	Potential wording	Origin of wording	Discussion with team
		Please mark your swollen joints:  Please mark your painful joints:  Please mark your stiff joints: 		With visual option - how do you quantify this? What image to use? E.g. with/out joints? How to mark e.g. shade (patient burden), tick (accuracy), mark (different for everyone). Feasibility is a big consideration here [JK]
Local and widespread: Location	Location	Capture the concept of all over stiffness e.g. In the past [insert timeframe] have you experienced all over/whole body stiffness?	QD	Could ask 'over the last 7 days I have experienced stiffness in 0 joints/1 joint/2 or 3 joints/all over [SH], 1 joint only/up to 5 joints/more than 5 joints, none/a few/many [SHa]. Any would capture severity [SH/JK] as might all over [SH] Patients really liked the idea about visual options [ED] but there are practical/feasibility considerations – see above. When run by GB - visual option would be do-able. Could provide people the option if that is easier for them rather than for everyone? [GB] Think about the specific wording for all over/whole body – check exact wording in transcripts [SH] As it is it is likely to need a Yes/No response. Don't place together as you could have one and not the other (whole body may be different from joints). Overall something capturing the number of joints, and something separate capturing all over stiffness) [JK/SH]
		Capture the concept of number of stiff joints e.g. In the past [insert timeframe] where have you felt stiff? One joint/two joints/etc	QD	
		I have experienced stiffness affecting more joints than usual?	QD, JK, SH	
		Capture the location of stiffness and whether that is 'all over' and 'usual' visually e.g. Please mark your stiff joints:  Please tick if the following are true for you: My whole body is stiff ☐ This is usual for me ☐	QD	
		Capture a pattern of responses - see McGill pain questionnaire e.g. Please tick all that have applied	QD, JK	

Key area	Specific aspect to capture	Potential wording	Origin of wording	Discussion with team
		to you in the past [insert timeframe]: all joints/ /many joints/some joints/no joints/outside the joints/my whole body		
Local and widespread: Structure	Structure	Capture stiffness in body/joints separately?	QD, JK, SH, ED	Need to be clearly separated from joints 'some people experience stiffness in their muscles or body in addition to or instead of joints. Thinking about your body…' [SH], or 'outside of the joints in the muscles or body' [JK], or 'not thinking about your joints…' [SH] These items would need to go next to each other to enhance understanding [JK]
		Include specific wording e.g. 'joint', 'all over', 'body' depending on the question. Could use PFW wording 'all over or in your joints' where relevant	QD, OS (PFQ)	
Linked to behaviour and environment: Movement	Immobility	Capture stiffness after a period of immobility e.g. Have you experienced stiffness after a period of immobility?	QD	I think stiffness after immobility is really interesting. I definitely think that thinking of the pathophysiological causes of stiffness potentially that this could distinguish different types of stiffness (different processes?) and so it's a good idea to keep it in for now [JK]. Think about wording - ensure that any item is clear that this is talking about immobility over longer than just sitting down for 5 minutes [JK]. Suggested wording 'some time/a period of time for example in a chair/in bed' [SH], 'after a period of immobility (for example sitting still for an hour) my stiffness is no different/a little bit worse/a lot worse/much worse' [JK]
	Movement	Capture whether stiffness reduces following exercise or movement e.g. During the past [insert timeframe] has your stiffness reduced following exercise or movement?	QD, JP	
	Restricted position	Do you get stiff as a result of being in a restricted position?	QD	
Linked to behaviour and environment: Medication effectiveness	Medication effectiveness	My current RA medications are having less of an effect on stiffness than usual/have not been managing stiffness in the way they usually do?	QD	Lots of qualitative discussion so if medications are the way that some patients assess stiffness then it could be important to test - I think this is something that I can imagine patients relating to, this is like a patient's way of viewing things. If that is the way that patients view things [JK] Is this different from self-management? Think about placement [SH/ED] Could look into something around 'usual medications having less of an effect' [SH]
		Please indicate the effectiveness of your current RA medications in relieving your stiffness	QD, OS (BASDAI)	
		Are your current RA medications relieving stiffness as they usually do?	QD	

Key area	Specific aspect to capture	Potential wording	Origin of wording	Discussion with team
Linked to behaviour and environment: Effect of lifestyle and environment	Weather	During the past [insert timeframe] has the weather affected your RA stiffness?	QD	Like medications, if these aspects are the ways that patients use to talk about and assess stiffness then they need to be considered in assessment e.g. characterising the nature of stiffness [JK]. What is the relevance clinically? [SH]. This is not about clinical usefulness, it's about characterising stiffness in a patient relevant way [JK]. Although there is work into barometric pressure [SH/JK]. Could test and see.
	Overdoing things	Capture stiffness after overdoing things	QD	
	Other	Capture any other aspects relevant to the patient but not listed?	QD	
Highly variable: Individual experience	Individual	No direct question but include individual wording e.g. 'to you' and/or 'usual for you'	QD, AE, GB	
Highly variable: Temporal pattern	Frequency	Capture whether stiffness occurs more often/frequently e.g. During the past [insert timeframe] has stiffness occurred more frequently?	QD	More frequently might be useful [SH]. I prefer does your morning stiffness come back more frequently during the day (24 hour period) than usual [PAM] Variability is different from temporal pattern e.g. 'my stiffness varies or comes and goes throughout the day [not at all/a little/a lot]' or 'my stiffness varies from day to day [not at all/a little/a lot]' [JK]. Temporal pattern is different from variability. Timing questions could be over the last 7 days e.g. 'have you had stiffness most mornings, most afternoons, most evenings, most nights' [JK]. The timing question/s are about severity - the more times you are stiff = more severity because if you are stiff for more times it means the total amount of time you are stiff is greater [JK] Same practical/feasibility considerations with visual options – see above
	Variability	Has stiffness been more changeable/variable than usual?	QD	
	When	Has stiffness affected you throughout the day and night?	QD	
	When	Capture when stiffness occurs rather than how long it lasts e.g. During the past [insert timeframe] when have you experienced RA stiffness? Night, morning, afternoon, evening	QD	
	When	Capture when stiffness occurs visually e.g. During the past [insert timeframe] when have you experienced RA stiffness? • Clock face • Chart (draw stiffness throughout the day?)	QD, JK, ED	
Highly variable: Duration	Duration	Capture whether stiffness has lasted longer than usual e.g. During the past [insert timeframe] has stiffness lasted longer than usual	QD, OS	This would capture aspects of fluctuation with disease.

Key area	Specific aspect to capture	Potential wording	Origin of wording	Discussion with team
		Capture duration through impact e.g. During the past [insert timeframe] has it taken you longer to complete your daily tasks because of stiffness?	QD	Asking about duration through impact could be ok [SH]. More about being slower [PR] which comes back to direct impact. Suggested wording 'stiffness makes me slower doing usual things' [PR/SH]
		Traditional stiffness items	OS	Important to include standard EMS duration items [JK/SH]
Highly variable: Severity	Severity	Capture whether stiffness has been worse than usual e.g. During the past [insert timeframe] has stiffness been worse than usual	QD, OS	This would capture aspects of fluctuation with disease.
		Capture severity through impact e.g. During the past [insert timeframe] have you found it difficult/been unable to complete your daily tasks because of stiffness?		
		Capture severity through aspects discussed above e.g. no. of affected joints? Frequency, variability		See discussion above
		Traditional stiffness items	OS	Include some standard severity questions [SH]
Impacts on daily life: Daily life impact	Sleep	Has stiffness made it difficult to sleep?/Has stiffness affected your sleep?	QD	Impact questions look sensible but there are a lot of them [JK]. Could be good one liners with similar responses to BRAF or other scale [SH]. The amount is disproportionate to the amount of other questions [JK]. Poor items will fall out if they are not good [SH]. Important to retain range of items highlighted in the qualitative work [GB]. Plus one impact question would not cover the whole of impact [GB] Layout needs to be considered - separate sections for clarity and to make a long list of questions look less intimidating [GB]. Layout must be appropriate for patients [SH] Think about the choice of wording to ensure it is accessible to all responders [ED]
	Dressing or undressing	Has stiffness made it difficult to dress and undress yourself?	QD	
	Washing	Have you found it difficult to bath or shower because of stiffness?	QD	
	Daily tasks or chores	Has stiffness made it difficult to do your usual daily tasks and activities?	QD	
	Daily tasks or chores: Eating	Has stiffness made it difficult to eat, including chewing and cutting food?	QD	
	Daily tasks or chores: Driving	Has stiffness made it difficult to drive a car/vehicle?	QD	

Key area	Specific aspect to capture	Potential wording	Origin of wording	Discussion with team
	Daily tasks or chores: Walking	Have you found it difficult to walk because of stiffness?	QD	
	Daily tasks or chores: Brush teeth	Has stiffness made it difficult brush your teeth?	QD	
	Work	Have you found it difficult to work because of stiffness?	QD	
	Hobbies	Have you found it difficult to do hobbies and activities you enjoy?	QD	
	Getting up	Have you found it difficult to get out of bed because of stiffness?	QD	
	Rise from chair	Have you found it difficult to rise from a chair because of stiffness?	QD	
	Effort	Have you found doing your usual daily tasks and activities requires more effort than usual?	QD	
	Impact	Stiffness has an impact on my daily life?	QD	
Impacts on daily life: Physical impact	Dexterity or fine movements	Has stiffness made it difficult to perform fine movements e.g. use a pen or pencil or do up buttons on a shirt or cardigan?	QD	
	Grip	Has stiffness made it difficult to grip or hold things e.g. opening a bottle of milk	QD	
	Grip: make a fist	Has stiffness made it difficult to make a fist?	QD	
	Loss of strength or weakness	Have you lacked strength to complete tasks or activities because of stiffness?	QD	
	ROM	Has your range of movement been reduced by stiffness?/Has stiffness limited your range of movement?	QD	
	Flexibility	Has stiffness limited your flexibility?	QD	

Key area	Specific aspect to capture	Potential wording	Origin of wording	Discussion with team
	Balance	Has stiffness made it difficult to balance?	QD	
	Slow	Has stiffness made you slow getting going/performing your usual daily tasks and activities?	QD	
	Slow: can't rush	Has stiffness made it difficult to rush?	QD	
	Mobility	Has stiffness made it difficult to move? Has stiffness restricted your mobility?	QD	
	Unbend limbs	Have you found you are unable to move parts of your body because of stiffness?	QD	
	Can't move	Have you found you are unable to move because of stiffness?	QD	
	Difficult to move	Has stiffness made it difficult to move?	QD	
	Can't move as expected	Has your body not moved like your brain expects it to because of stiffness?	QD	
Impacts on daily life: Cognitive impact	Can't move as expected	Has your body not moved like your brain expects it to because of stiffness?	QD	
	Thinking: concentrate	Have you had to concentrate to move perform tasks because of stiffness?	QD	
	Thinking: think differently	Have you had to think differently because of stiffness?	QD	
Impacts on daily life: Psychosocial impact	Frustration	Have you felt frustrated because of stiffness?	QD	
	Worry	Have you felt worried because of stiffness?	QD	
	Embarrassed	Have you felt embarrassed because of stiffness?	QD	
Impacts on daily life: Pain	Pain on movement	Has stiffness made it difficult to move without pain?	QD	See arguments about regarding including pain directly e.g. retain focus on stiffness [GB], confusing from measurement perspective [JP], pain captured in validity testing already
		Have you been able to self-manage stiffness?	QD	

Key area	Specific aspect to capture	Potential wording	Origin of wording	Discussion with team
Requires self-management: Direct	Self-management	Have you had to spend more time than usual managing stiffness?	QD	From the HIT day people said that they don't use the word 'self-manage' [ED]. That is also true of the qualitative data – 'self-management' was not the term used [SHa]. 'Cope' or 'deal' is more of a patient word [PR], or 'made do', or 'work around' but that has more of a problem solving or practical slant. Think about the options [SH/ED] Think about the choice of wording to ensure it is accessible to all responders [SH] Could look at having to use more strategies for example medications/hot cold as it might be a level of severity issue [SH] Look at heath foundation for further information on measuring self-management [ED]
		Capture self-management in the impact triad	QD	
		Work around stiffness	QD	
Requires self-management: Indirect		Have you had to change your plans or behaviour because of stiffness?	QD	
		Have you had to accept help from others because of stiffness?	QD	
		Change plans or behaviour?	QD	
		Have you been able to self-manage stiffness as well as usual?	QD	
Key: QD = qualitative data, initials (JK, JP, ED, SH, GB, AE, PR) = discussion with supervisory team member, OS = other scale				

Appendix P: Item tracking matrix

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
1	Severity	Local and widespread: Location	Location	Thinking about the location of your RA stiffness: How many joints have you experienced RA stiffness in during the last 7 days?	4 option ordinal scale (none of my joints, some of my joints, many of my joints, all of my joints)	Location is not the right work [SH] How many indicates countable quality [ED]	Have you experienced RA stiffness in your joints during the past 7 days?	4 option ordinal scale (not in any of my joints, yes, in some of my joints, yes, in many of my joints, yes, in all of my joints)
2	Severity	Local and widespread: Location	Location	Thinking about the location of your RA stiffness: How many parts of the body have you experienced RA stiffness in during the last 7 days?	4 option ordinal scale (no parts of my body, some parts of my body, many parts of my body, my whole body)	How many indicates countable quality [ED] Clarify that this is about stiffness in your body but outside the joints [ED] Is the 'whole body' is different to 'parts of the body' [ED]	Have you experienced RA stiffness in your body (outside of your joints) during the past 7 days?	4 option ordinal scale (not in any parts of my body, yes, in some parts of my body, yes, in many parts of my body, yes, in my whole body)
3	Severity	Highly variable: Temporal pattern	Variability	During the last 7 days my RA stiffness has varied (comes and goes) throughout the day or night/from day to day?	4 option ordinal scale (not at all, a little, quite a bit, very much)	Varied in intensity or in number of episodes? [ED] Could clarify by asking whether stiffness has occurred more frequently than usual? [ED]	During the past 7 days have you experienced RA stiffness coming and going as frequently as usual for you?	5 option Likert scale (it has been much less frequent than usual, it has been less frequent than usual, it has been the same as usual, it has been more frequent than usual,

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
						Day or night sounds like a choice – could use 'throughout the day' or maybe just get rid of that bit and end with throughout [ED]		it has been much more frequent than usual)
4	Severity	Linked to behaviour and environment: Movement	Immobility	After a period of immobility (e.g. in a chair or in bed) my stiffness is	4 option ordinal scale (no different, a little bit worse, quite a bit worse, much worse)		During the past 7 days have you experienced stiffness after a period of immobility (for example, in a chair or in bed)?	5 option Likert scale (I have had much less than usual, I have had less than usual, I have had the same as usual, I have had more than usual, I have had much more than usual)
5	Severity	Linked to behaviour and environment: Medication effectiveness	Medication effectiveness	My RA medications have been affecting my RA stiffness	5 option Likert scale (much less well than usual, less than usual, the same as usual, better than usual, much better than usual)		During the past 7 days have your RA medications been controlling RA stiffness as usual for you?	5 option Likert scale (it has been much less well controlled than usual, it has been less well controlled than usual, it has been controlled the same as usual, it has been controlled better than usual, it has been controlled

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
								much better than usual)
6	Severity	Highly variable: Temporal pattern	When	During the last 7 days when have you experienced RA stiffness?	4 option ordinal scale (During the morning, During the afternoon, During the evening, During the night)	Ensure you address somewhere that there are two very similar items – either verbally or written [JK]	During the past 7 days when have you experienced RA stiffness (in your joints or your body)?	4 option ordinal scale (During the morning, During the afternoon, During the evening, During the night)
7	Severity	Highly variable: Temporal pattern	When	Thinking about your RA stiffness during the last 7 days, please mark on the clock face the times of day when you have experienced RA stiffness	Clock face (visual option)		During the past 7 days when have you experienced RA stiffness (in your joints or your body)?	8 option ordinal scale (first thing when I wake up, when I get out of bed, during the first few hours after I get up, during the late morning, during the early afternoon, during the late afternoon, during the evening, during the night)
8	Impact	Impacts on daily life: Daily life impact	Sleep	Has stiffness affected your sleep?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness affected your sleep?	4 option ordinal scale (not at all, a little, quite a bit, very much)
9	Impact	Impacts on daily life: Daily life impact	Dress/ undress	Has stiffness made it difficult to dress or undress yourself?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to dress or undress yourself?	4 option ordinal scale (not at all, a little, quite a bit, very much)
10	Impact	Impacts on daily life: Daily life impact	Wash	Has stiffness made it difficult to bath or shower?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to bath or shower?	4 option ordinal scale (not at all, a little, quite a bit, very much)

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
11	Impact	Impacts on daily life: Daily life impact	Daily tasks/ chores/work	Has stiffness made it difficult to work or do other daily activities?	4 option ordinal scale (not at all, a little, quite a bit, very much)	Would work need a does not apply option? [SH] Would none not capture does not apply? [ED]	Has stiffness made it difficult to work?	4 option ordinal scale (not at all, a little, quite a bit, very much)
12	Impact	Impacts on daily life: Daily life impact	Daily tasks/ chores/work	Has stiffness made it difficult to work or do other daily activities?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to do your daily activities?	4 option ordinal scale (not at all, a little, quite a bit, very much)
13	Impact	Impacts on daily life: Daily life impact	Eat/cut food	Has stiffness made it difficult to eat? e.g. chewing or cutting your food	4 option ordinal scale (not at all, a little, quite a bit, very much)	If not trying to save space use 'for example' rather than e.g. [JK]	Has stiffness made it difficult to eat? For example, chew or cut your food	4 option ordinal scale (not at all, a little, quite a bit, very much)
14	Impact	Impacts on daily life: Daily life impact	Hobbies	Has stiffness made it difficult to do hobbies or activities you enjoy?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to do hobbies or activities you enjoy?	4 option ordinal scale (not at all, a little, quite a bit, very much)
15	Impact	Impacts on daily life: Daily life impact	Get out of bed	Has stiffness made it difficult to get out of bed?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to get out of bed?	4 option ordinal scale (not at all, a little, quite a bit, very much)
16	Impact	Impacts on daily life: Daily life impact	Rise from a chair	Has stiffness made it difficult to rise from a chair?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to rise from a chair?	4 option ordinal scale (not at all, a little, quite a bit, very much)
17	Impact	Impacts on daily life: Daily life impact	Effort	Do your daily activities require more effort than usual because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Have your daily activities required more effort than usual because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
18	Impact	Impacts on daily life: Daily life impact	Overall impact	Has stiffness had an impact/effect on your daily life?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness had an impact on your daily life?	4 option ordinal scale (not at all, a little, quite a bit, very much)
19	Impact	Impacts on daily life: Physical impact	Slow	Has stiffness made you slower doing your daily activities?	4 option ordinal scale (not at all, a little, quite a bit, very much)	Could be misinterpreted – think about wording [SH]	Has stiffness made you slower? For example, unable to rush	4 option ordinal scale (not at all, a little, quite a bit, very much)
20	Impact	Impacts on daily life: Physical impact	Dexterity	Has stiffness made it difficult to do fine movements? e.g. Do up buttons on a shirt or cardigan?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to do fine movements? For example, do up buttons on a shirt or cardigan?	4 option ordinal scale (not at all, a little, quite a bit, very much)
21	Impact	Impacts on daily life: Physical impact	Grip	Has stiffness made it difficult to grip or hold things?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to grip or hold things?	4 option ordinal scale (not at all, a little, quite a bit, very much)
22	Impact	Impacts on daily life: Physical impact	Make a fist	Has stiffness made it difficult to make a fist?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to make a fist?	4 option ordinal scale (not at all, a little, quite a bit, very much)
23	Impact	Impacts on daily life: Physical impact	Strength	Have you lacked strength to do your daily activities because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Have you lacked physical strength to do your daily activities because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)
24	Impact	Impacts on daily life: Physical impact	Flexibility/ ROM/ restricted movement	Has stiffness reduced your range of movement/ flexibility?	4 option ordinal scale (not at all, a little, quite a bit, very much)	What do you mean by flexibility? Go back to patient data [SH]	Have you found that your movement is restricted because of stiffness? For	4 option ordinal scale (not at all, a little, quite a bit, very much)

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
							example, reaching to get an item	
25	Impact	Impacts on daily life: Physical impact	Balance	Has stiffness reduced your balance?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to balance without physically supporting yourself?	4 option ordinal scale (not at all, a little, quite a bit, very much)
26	Impact	Impacts on daily life: Physical impact	Difficult to move	Has stiffness made it difficult to move? (Parts of your body or your whole body)	4 option ordinal scale (not at all, a little, quite a bit, very much)		Has stiffness made it difficult to move parts of your body or your whole body?	4 option ordinal scale (not at all, a little, quite a bit, very much)
27	Impact	Impacts on daily life: Cognitive impact	Don't move as expected	Has your body not moved like your brain expects it to because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)	Unsure about 'expects', could change to 'tells' [SH]	Has your body not moved like your brain tells it to because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)
28	Impact	Impacts on daily life: Cognitive impact	Concentrate	Have you had to concentrate more than usual to move your body because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Have you had to concentrate more than usual to move your body because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)
29	Impact	Impacts on daily life: Psychosocial impact	Frustration	Have you felt frustrated because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Have you felt frustrated because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)
30	Impact	Impacts on daily life: Psychosocial impact	Worry	Have you felt worried because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Have you felt worried because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)
31	Impact	Impacts on daily life:	Embarrassed	Have you felt embarrassed	4 option ordinal scale (not at all, a	Consider that it is difficult to translate embarrassed [JK]	Have you felt embarrassed	4 option ordinal scale (not at all, a

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
		Psychosocial impact		because of stiffness?	little, quite a bit, very much)		because of stiffness?	little, quite a bit, very much)
32	Impact/ duration	Highly variable: Duration	Duration through impact	Has it taken you longer to complete your daily activities because of RA stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)	Could change 'complete' to something more simple, maybe 'do' [ED]	Has it taken you longer to do your daily activities because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)
33	Impact/ severity	Highly variable: Severity	Severity through impact	Have you been unable to complete your daily activities because of RA stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)	Could change 'complete' to something more simple, maybe 'do' [ED]	Have you been unable to do your daily activities because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)
34	Impact	Requires self-management: Indirect	Indirect self-management	Have you had to change your plans or behaviour because of RA stiffness?			Have you had to change your plans or behaviour because of stiffness	4 option ordinal scale (not at all, a little, quite a bit, very much)
35	Impact	Requires self-management: Direct	Direct self-management				Have you had to work around stiffness more than usual?	4 option ordinal scale (not at all, a little, quite a bit, very much)
36	Impact	Requires self-management: Indirect	Indirect self-management	Have you had to accept help from others because of RA stiffness?			Have you had to accept help from others because of stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)
37	Impact	Requires self-management: Direct	Direct self-management	Have you had to spend more time than usual dealing with stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)		Have you had to spend more time than usual coping (managing, dealing with, making do) with stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
38	Impact	Requires self-management: Direct	Direct self-management	Thinking about how well you have coped with your RA stiffness during the last 7 days: I have been able to deal with RA stiffness	5 option Likert scale (much less well than usual, less well than usual, same as usual, better than usual, much better than usual)		Have you been able to cope (manage, deal, make do) with stiffness?	4 option ordinal scale (not at all, a little, quite a bit, very much)
39	Impact	Impacts on daily life	Impact triad – severity	Please circle the number/mark the line/tick the box that best describes your level of RA stiffness (in your joints or your whole body, and at any time of day) during the past 7 days	NRS (0 = no stiffness, 10 = extreme stiffness)	Think about the response options for these items [SH] Could remove 'whole' here for consistency with item 2 [ED] Uncertainty about the response options for the ordinal response options – unsure about the use of quotes especially for the self-management item as this will be so different for everybody [ED] Uncertainty about the response options for the	Please circle the number/mark the line/tick the box that best describes the severity of your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days	NRS (0 = no stiffness, 10 = extreme stiffness)
46					NRS (0 = no stiffness, 10 = extreme stiffness)			VAS (0 = no stiffness, 10 = extreme stiffness)
53					4 ordinal options (no stiffness "I have not been stiff", some stiffness "I have been a bit stiff in the morning and/or have seized up a bit during the day", quite a lot of stiffness "I have been quite stiff in the morning and/or have seized up quite a lot during the day", extreme stiffness "I have been so stiff I have been unable to move at all")			4 ordinal options (no stiffness, a little stiffness, quite a bit of stiffness, extreme stiffness)

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
40	Impact	Impacts on daily life	Impact triad – self-management	Please circle the number/mark the line/tick the box that best describes how well you have been able to deal with your RA stiffness (in your joints or your whole body, and at any time of day) during the past 7 days	NRS (0 = not well at all, 10 = very well)	ordinal response options – think about whether each option is a logical step up from the one before [ED] Initially used 'last 7 days' then changed to 'during the past 7 days' or 'over the past 7 days' – decide which to use and keep consistent [JK] Change the introduction to the impact items to something more concise [JK] Level – flat on the ground?? [JK] Think about wording of self-management items – why using all these different options? Include if there is a reason to as it is an extra	Please circle the number/mark the line/tick the box that best describes how well you have been able to deal with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days	NRS (0 = not well at all, 10 = very well)
47					VAS (0 = not well at all, 10 = very well)			VAS (0 = not well at all, 10 = very well)
54					4 ordinal options (I have not had any stiffness to deal with, I have been able to deal with my stiffness by using ways to work around it, I have been able to deal with some of my stiffness but I have also had to change my plans and ask for help because of it, I have been unable to deal with my stiffness)			4 ordinal options (not at all well, not very well, quite well, very well)
41	Impact	Impacts on daily life	Impact triad – self-management	Please circle the number/mark the line/tick the box that best describes how well you have been able to cope with your RA stiffness (in your joints or your whole body, and at any time of	NRS (0 = not well at all, 10 = very well)		Please circle the number/mark the line/tick the box that best describes how well you have coped with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days	NRS (0 = not well at all, 10 = very well)
48					VAS (0 = not well at all, 10 = very well)			VAS (0 = not well at all, 10 = very well)
55					4 ordinal options (I have not had any stiffness to cope with, I have been able to deal with my stiffness by using			4 ordinal options (not at all well, not very well, quite well, very well)

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
				day) during the past 7 days	ways to work around it, I have been able to cope with some of my stiffness but I have also had to change my plans and ask for help because of it, I have been unable to cope with my stiffness)	burden on the participant [JK]		
42	Impact	Impacts on daily life	Impact triad – self-management	Please circle the number/mark the line/tick the box that best describes how well you have been able to <u>manage</u> with your RA stiffness (in your joints or your whole body, and at any time of day) during the past 7 days	NRS (0 = not well at all, 10 = very well)		Please circle the number/mark the line/tick the box that best describes how well you have <u>managed</u> your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days	NRS (0 = not well at all, 10 = very well)
49					VAS (0 = not well at all, 10 = very well)			VAS (0 = not well at all, 10 = very well)
56					4 ordinal options (I have not had any stiffness to manage, I have been able to manage my stiffness by using ways to work around it, I have been able to manage some of my stiffness but I have also had to change my plans and ask for help because of it, I have been unable			4 ordinal options (not at all well, not very well, quite well, very well)

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
					to manage my stiffness)			
43	Impact	Impacts on daily life	Impact triad – self-management	Please circle the number/mark the line/tick the box that best describes how well you have been able to <u>coped (managed, dealt with, made do)</u> with your RA stiffness (in your joints or your whole body, and at any time of day) during the past 7 days	NRS (0 = not well at all, 10 = very well)		Please circle the number/mark the line/tick the box that best describes how well you have <u>coped (managed, dealt with, made do)</u> with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days	NRS (0 = not well at all, 10 = very well)
50					VAS (0 = not well at all, 10 = very well)			VAS (0 = not well at all, 10 = very well)
57					4 ordinal options (I have not had any stiffness to cope (manage, deal with, make do) with, I have been able to deal with my stiffness by using ways to work around it, I have been able to cope (manage, deal with, make do) with some of my stiffness but I have also had to change my plans and ask for help because of it, I have been unable to cope (manage, deal with, make do) with my stiffness)			4 ordinal options (not at all well, not very well, quite well, very well)
44	Impact	Impacts on daily life	Impact triad – importance	Please circle the number/mark the	NRS (0 = no effect, 10 = great effect)		Please circle the number/mark the	NRS (0 = no effect, 10 = great effect)

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
51				line/tick the box in the position that best describes the effect RA stiffness (in your joints or your whole body, and at any time of day) has had on your life during the past 7 days	VAS (0 = no effect, 10 = great effect)		line/tick the box that best describes the effect RA stiffness (in your joints or your body, and at any time of day) has had on your life during the past 7 days	VAS (0 = no effect, 10 = great effect)
58					4 ordinal options (no effect at all, a little effect because although it has been there I have been able to get on with life, quite a bit of an effect because it has restricted me doing my daily activities, a great effect because I have been so stiff I have been unable to do my usual daily activities)			4 ordinal options (no effect, a little effect, quite a bit of an effect, a great effect)
45	Impact	Impacts on daily life	Impact triad – impact	Please circle the number/mark the line/tick the box that best describes the impact of RA stiffness (in your joints or your whole body, and at any time of day) during the past 7 days	NRS (0 = no impact, 10 = a great impact)		Please circle the number/mark the line/tick the box that best describes the overall impact on your life of RA stiffness (in your joints or your body, and at any time of day) during the past 7 days	NRS (0 = no impact, 10 = a great impact)
52					VAS (0 = no impact, 10 = a great impact)			VAS (0 = no impact, 10 = a great impact)
59					4 ordinal options (no impact, a little impact, quite a bit of an impact, a great impact)			4 ordinal options (no impact, a little impact, quite a bit of an impact, a great impact)
60	Attribution	Linked to behaviour and environment:	Weather	Thinking about the nature of your RA stiffness over	4 option ordinal scale (not at all, a	Think about the wording and	Please indicate how much the weather has	4 option ordinal scale (No contribution, some

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
		Effect of lifestyle and environment		[insert timescale] My RA stiffness is affected by the weather	little, quite a bit, very much)	response options for these items [SH] What is this trying to get at – cause or contribution as cause sounds factual. Also this is asking about peoples experiences and their perceptions so important to highlight that e.g. 'how much of the following has contributed to your experience of …' [ED] The response options don't match the questions – think about these more. Maybe 'no contribution, a small contribution, a moderate contribution, the main contribution' [JK]	contributed to your experience of RA stiffness during the last 7 days	contribution, moderate contribution, the main contribution)
61	Attribution	Linked to behaviour and environment: Movement	Immobility	Thinking about the nature of your RA stiffness over [insert timescale]. My RA stiffness is affected by being immobile for a period of time (e.g. in a chair or in bed)	4 option ordinal scale (not at all, a little, quite a bit, very much)		Please indicate how much being immobile for a period of time for example, in a chair or in bed has contributed to your experience of RA stiffness during the last 7 days	4 option ordinal scale (No contribution, some contribution, moderate contribution, the main contribution)
62	Attribution	Part of having RA: Relationship with other symptoms	Inflammation	Thinking about the nature of your RA stiffness over [insert timescale] My RA stiffness is affected by other RA symptoms (e.g. swelling or inflammation)	4 option ordinal scale (not at all, a little, quite a bit, very much)		Please indicate how much inflammation or swelling have contributed to your experience of RA stiffness during the last 7 days	4 option ordinal scale (No contribution, some contribution, moderate contribution, the main contribution)
63	Attribution	Part of having RA: Process	Process	Thinking about the nature of your RA stiffness over [insert timescale] My RA stiffness is affected by being in a flare	4 option ordinal scale (not at all, a little, quite a bit, very much)		Please indicate how much being in an RA flare (flare up) has contributed to your experience of RA stiffness during the last 7 days	4 option ordinal scale (No contribution, some contribution, moderate contribution, the main contribution)
64	Attribution	Part of having RA: Process	Process	Thinking about the nature of your RA	4 option ordinal scale (not at all, a		Please indicate how much joint	4 option ordinal scale (No

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
				stiffness over [insert timescale] My RA stiffness is affected by joint damage	little, quite a bit, very much)	Specify what you are looking for in item 68 e.g. 'please specify anything that applies to you but is not listed above' [JK]	damage has contributed to your experience of RA stiffness during the last 7 days	contribution, some contribution, moderate contribution, the main contribution)
65	Attribution	Part of having RA: Relationship with other symptoms	Pain	Thinking about the nature of your RA stiffness over [insert timescale] My RA stiffness is affected by other RA symptoms (e.g. pain)	4 option ordinal scale (not at all, a little, quite a bit, very much)		Please indicate how much other RA symptoms have contributed to your experience of RA stiffness during the last 7 days	4 option ordinal scale (No contribution, some contribution, moderate contribution, the main contribution)
66	Attribution	Linked to behaviour and environment: Effect of lifestyle and environment	Overdoing things	Thinking about the nature of your RA stiffness over [insert timescale] My RA stiffness is affected by overdoing things	4 option ordinal scale (not at all, a little, quite a bit, very much)		Please indicate how much overdoing things for example, doing too much has contributed to your experience of RA stiffness during the last 7 days	4 option ordinal scale (No contribution, some contribution, moderate contribution, the main contribution)
67	Attribution	Linked to behaviour and environment: Medication effectiveness	Medication effectiveness	My RA medications have been affecting my RA stiffness	5 option Likert scale (much less well than usual, less than usual, the same as usual, better than usual, much better than usual)		Please indicate how much RA medications not controlling your disease has contributed to your experience of RA stiffness during the last 7 days	4 option ordinal scale (No contribution, some contribution, moderate contribution, the main contribution)
68	Attribution	Linked to behaviour and	Other	Thinking about the nature of your RA	4 option ordinal scale (not at all, a		Please indicate how much [please	4 option ordinal scale (No

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
		environment: Effect of lifestyle and environment		stiffness over [insert timescale] My RA stiffness is affected by [please insert]	little, quite a bit, very much)		specify] anything that applies to you but is not listed above has contributed to your experience of RA stiffness during the last 7 days	contribution, some contribution, moderate contribution, the main contribution)
69	Traditional	Highly variable: Severity	Severity	NA	NA	NA	How would you describe the overall level of morning stiffness you have had from the time you wake up?	VAS (0 = No stiffness 10 = Very severe stiffness)
70								NRS (0 = No stiffness 10 = Very severe stiffness)
71								5 ordinal options (no stiffness, mild stiffness, moderate stiffness, severe stiffness, very severe stiffness)
72	Traditional	Highly variable: Duration	Duration	NA	NA	NA	How long does your morning stiffness last from waking until maximum improvement occurs?	Minutes or hours
74								3 ordinal options (up to an hour, 1 – 3 hours, ≥3 hours)
73	Traditional	Highly variable: Duration	Duration	NA	NA	NA	Were your joints stiff when you woke up today? If yes, how long did this stiffness last?	6 ordinal options (less than 30 minutes, 30 minutes–1 hour, 1–2 hours, 2–4 hours, over 4 hours, all day)

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
75	Response shift	Part of having RA: Normal and specific nature	Normality	Thinking about a usual week when you are not in a flare of your RA. Please circle the number that shows your usual level of RA stiffness	NRS (0 = no stiffness, 10 = extreme stiffness)	Could test with different formats [SH]	Thinking about a usual week when you are not in a flare (flare-up) of your RA please [mark the response] that shows your usual RA stiffness (in your joints or your body, and at any time of day)	NRS (0 = no stiffness, 10 = extreme stiffness)
76								VAS (0 = no stiffness, 10 = extreme stiffness)
77								4 option ordinal scale (no stiffness, some stiffness, quite a bit of stiffness, extreme stiffness)
-	Intro-duction	Part of having RA: Process	Process	"[…] Although we acknowledge that some people with RA have joints that are always difficult to move whether their RA is good or bad (due to joint damage), please do not include this when you answer this questionnaire. This questionnaire will help us understand your disease activity. We will use the answers that you give alongside other tests (e.g. blood tests) to build	NA	This is pretty wordy – worth trying to make it as concise as possible [SH] Shorter is better [ED] The reason for doing the questionnaire will be captured in the context anyway [ED] If using underlining you need to underline the whole relevant part – change and ensure consistent [JK]	"[…] It will help us understand how active your disease is. Some people have joints that are always difficult to move whether their RA is good or bad (for example, due to joint damage). Please do not include this sort of stiffness when you answer this questionnaire"	NA

Item no.	Item section	Key area	Specific aspect	Original wording	Original response options	Changes suggested by team	Wording taken to Study 3	Response options taken to Study 3
				a picture of your disease"				
-	Wording	Part of having RA: Normal and specific nature	Specificity	No specific wording but 'stiffness' or 'RA stiffness' used in stem questions	NA	Important to reiterate stiffness in every stem question [SH] Ensure consistent throughout [JK]	No specific wording but 'RA stiffness' used in every stem question	NA
-	Wording	Local and widespread: Structure	Structure	No specific wording but use relevant wording e.g. 'joint' or 'all over or in your joints'	NA			NA
-	Wording	Highly variable: Individual experience	Individual	No specific wording but using 'to you' or other individual specific wording	NA		No specific wording but using 'to you' or other individual specific wording	NA

Appendix Q: Patient information sheet (Study 3)

A research study to explore stiffness in people with rheumatoid arthritis (Phase 2)

Patient information sheet

You are being invited to take part in a research study. Before you decide whether to take part, it is important for you to understand why the research is being done, and what it will involve. Please take time to read the following information leaflet carefully and discuss it with others if you wish. If anything is not clear or you would like more information please ask one of the team.

What is the purpose of the study?
One of the problems commonly experienced by people with rheumatoid arthritis is stiffness. Stiffness is a term used by patients and health professionals, but it is not very well understood. The aim of this study is to develop a way of measuring stiffness in rheumatoid arthritis. With the help of other people like yourself we have developed a better understanding of what stiffness means to people. We would now like your help to develop this further by working towards a way of measuring stiffness in a questionnaire.

Why have I been chosen?
You have been chosen because you have rheumatoid arthritis.

Who is asking me to take part?
I am Serena Halls, a PhD student at the University of the West of England. This research study is the second of three studies which will form part of my PhD.

What will I be asked to do if I take part?
You will be asked to attend an informal interview with the researcher (Serena Halls) at the rheumatology department you attend. Before the interview starts, the researcher will ask you to read and sign the consent form and ask you some questions about your medical history. During the interview, she will invite you to answer a series of questions that have been designed to measure stiffness in rheumatoid arthritis. She will encourage you to 'think aloud' as you answer the questions and will prompt you about any thoughts or preferences you might have about the questions. You can say as much or as little as you like, there are no right or wrong answers. The interview will last for about an hour and we will offer you refreshments and are happy to pay your travel costs. We will ask your permission to audio-record the interview, which we will type up (transcribe) and then analyse after the interview.

Do I have to take part?
No, taking part is voluntary. It is up to you to decide whether or not to take part. If you decide not to take part you do not need to give a reason, nobody will be upset and the standard of care you receive will not be affected. If you decide to take part we will ask you to sign a consent form, and will give you a copy of this information sheet and the consent form to keep. For general advice about taking part in research, you can contact the local Patient Advice and Liaison Service on 0117 900 3433 or pals@bristolpct.nhs.uk.

What if I wish to withdraw at a later stage?
You are free to withdraw from the study at any time, and with no explanation.

What are the risks of taking part in the study?
We do not believe there are any risks in being involved in this study. We appreciate that there may be some inconvenience to you by having to come into the hospital for the interview but we will try and reduce this by arranging a convenient date and time for you to come.

What are the benefits of taking part in the study?
The benefits of taking part in this study are that you will be helping us to gain a better understanding of how to measure stiffness in people with rheumatoid arthritis. This will help us to improve decisions made about treatment and management of rheumatoid arthritis.

Will my taking part in the study be kept confidential?
Yes. Your name will be replaced by a code. All other identifying information (such as people's names, locations or specific descriptions) will be replaced with code numbers or a generalised summary. No one will be able to identify you from any analysis or report. The study reports will include quotations from the interviews but no names will be used. The recordings will be kept securely for 6 years and then destroyed, in accordance with best practice in research guidelines.

What will happen to the results of this research study?
Research team members will analyse the anonymous transcripts and discuss our findings. The findings of this study will influence the design of later research studies within this PhD. We hope the results will be reported in professional journals and at meetings (but participants will not be identified by name). We will send you a summary of the results if you would like.

Who is organising, funding and reviewing the research?
The study is coordinated by a team from the university of the West of England (UWE) based at the Academic Rheumatology Unit at the Bristol Royal Infirmary. It is funded by UWE and has been peer reviewed by the local and UWE Research Ethics Committees.

What do I do now?
Thank you for considering taking part in this research. Please return the reply slip provided if you would like to take part by returning it in the pre-paid reply envelope to Serena Halls. Serena will then contact you in a few days with further information and to answer any questions.

Research team:

Serena Halls, PhD Student Researcher, UWE Bristol
Professor Sarah Hewlett, Professor of Rheumatology Nursing, UWE Bristol
Professor John Kirwan, Professor of Rheumatic Diseases, UoB Bristol
Dr Jon Pollock, Reader in Epidemiology, UWE Bristol
Dr Emma Dures, Research Fellow, UWE Bristol
Mrs Avis Edmunds, Patient Research Partner
Mrs Gill Baker, Patient Research Partner

Contact:

Serena Halls	0117 342 4972	Serena.Halls@uwe.ac.uk
Sarah Hewlett	0117 324 2903	Sarah.Hewlett@uwe.ac.uk

Appendix R: Cognitive interview topic guide (Study 3)

Study ID: _____ Date: _____ Start time: _____

Instructions for interviewee:

Note to the interviewer:

a) *Either read these instructions in their entirety or paraphrase them*
b) *Note that this form is set up to be administered after the interviewee has signed the consent form*

Thank you for coming in. Let me tell you a little more about what we will be doing today

- We are not collecting information about you but are trying our questions out on people like you so we can make the questions better
- Our goal here is to get a better idea of how the questions are working. So I'd like you to *think aloud* as you answer the questions-just tell me *everything* that comes to mind as you go about answering them
- At times I might ask you about what you think a question is asking about, how you come up with your answers and how you interpret the questions
- Some of the questions might look very similar. This is because we are trying to find the best way to word the question, so if there are things you particularly like or don't like please do say!
- If any question is unclear, hard to answer or doesn't make sense please tell me that – don't be shy! There are no right or wrong answers.
- Finally, we will do this for about an hour unless I run out of things to ask you before then
- Do you have any questions before we start?

(Adapted from Willis, 2005)

TURN TAPE RECORDER ON

Refer to questionnaire: Questions are written on other sheet with space for notes

General probes:
In your own words, what is this question asking?
How did you come up with that answer?
Tell me more about that
What does the term...mean to you in this question?
What time period are you thinking of?
You hesitated a bit there, what are you wondering about?
How could we phrase that question better?

End time of interview: _____

Appendix S: Draft PROM items for cognitive interviews (Study 3)

Rheumatoid Arthritis Stiffness Questionnaire
(RAST-Q)

This questionnaire is about stiffness related to your rheumatoid arthritis or **RA stiffness**. It will help us understand **how active your disease is**. Some people have joints that are always difficult to move whether their RA is good or bad (for example, due to joint damage). Please do not include this sort of stiffness when you answer this questionnaire.

We would like to know how **RA stiffness** has affected you during the **past 7 days**

For each of the following questions, please tick <u>one</u> answer that best applies to you

1. Have you experienced RA stiffness **<u>in your joints</u>** during the past 7 days?

   - Not in any of my joints ☐
   - Yes, in some of my joints ☐
   - Yes, in many of my joints ☐
   - Yes, in all of my joints ☐

2. Have you experienced RA stiffness **<u>in your body</u>** (outside of your joints) during the past 7 days?

   - Not in any of my body ☐
   - Yes, in some parts of my body ☐
   - Yes, in many parts of my body ☐
   - Yes, in my whole body ☐

3. During the past 7 days have you experienced RA stiffness coming and going as frequently as usual for you?

   - It has been **<u>much less</u>** frequent than usual ☐
   - It has been **<u>less</u>** frequent than usual ☐
   - It has been the **<u>same</u>** as usual ☐
   - It has been **<u>more</u>** frequent than usual ☐
   - It has been **<u>much more</u>** frequent than usual ☐

4. During the past 7 days have you experienced stiffness after a period of immobility (for example, in a chair or in bed)?

   - I have had **<u>much less</u>** than usual ☐
   - I have had **<u>less</u>** than usual ☐
   - I have had the **<u>same</u>** as usual ☐
   - I have had **<u>more</u>** than usual ☐
   - I have had **<u>much more</u>** than usual ☐

5. During the past 7 days have your RA medications been controlling RA stiffness as usual for you?

It has been **much less well** controlled than usual	☐
It has been **less** well controlled than usual	☐
It has been controlled the **same** as usual	☐
It has been **better** controlled than usual	☐
It has been **much better** controlled than usual	☐

For each of the following questions, please tick all answers that apply to you

6. During the past 7 days **when** have you experienced RA stiffness (in your joints or your body)?

During the morning	☐
During the afternoon	☐
During the evening	☐
During the night	☐

7. During the past 7 days **when** have you experienced RA stiffness (in your joints or your body)?

First thing when I wake up	☐
When I get out of bed	☐
During the first few hours after I get up	☐
During the late morning	☐
During the early afternoon	☐
During the late afternoon	☐
During the evening	☐
During the night	☐

For each of the following questions, please tick <u>one</u> answer that best applies to you

Thinking about your RA stiffness during the **<u>past 7 days</u>**	**Not at all**	**A little**	**Quite a bit**	**Very much**	**Does not apply to me**
8. Has stiffness affected your sleep?	__	__	__	__	
9. Has stiffness made it difficult to dress or undress yourself?	__	__	__	__	
10. Has stiffness made it difficult to bath or shower?	__	__	__	__	
11. Has stiffness made it difficult to work?	__	__	__	__	__
12. Has stiffness made it difficult to do your daily activities?	__	__	__	__	
13. Has stiffness made it difficult to eat? For example, chew or cut your food	__	__	__	__	
14. Has stiffness made it difficult to do hobbies or activities you enjoy?	__	__	__	__	
15. Has stiffness made it difficult to get out of bed?	__	__	__	__	
16. Has stiffness made it difficult to rise from a chair?	__	__	__	__	
17. Have your daily activities required more effort than usual because of stiffness?	__	__	__	__	
18. Has stiffness had an impact on your daily life?	__	__	__	__	
19. Has stiffness made you slower? For example, unable to rush	__	__	__	__	
20. Has stiffness made it difficult to do fine movements? For example, do up buttons on a shirt or cardigan?	__	__	__	__	
21. Has stiffness made it difficult to grip or hold things?	__	__	__	__	
22. Has stiffness made it difficult to make a fist?	__	__	__	__	

Thinking about your RA stiffness during the **past 7 days**	**Not at all**	**A little**	**Quite a bit**	**Very much**	**Does not apply to me**
23. Have you lacked physical strength to do your daily activities because of stiffness?	__	__	__	__	
24. Have you found that your movement is restricted because of stiffness? For example, reaching to get an item	__	__	__	__	
25. Has stiffness made it difficult to balance without physically supporting yourself?	__	__	__	__	
26. Has stiffness made it difficult to move parts of your body or your whole body?	__	__	__	__	
27. Has your body not moved like your brain tells it to because of stiffness?	__	__	__	__	
28. Have you had to concentrate more than usual to move your body because of stiffness?	__	__	__	__	
29. Have you felt frustrated because of stiffness?	__	__	__	__	
30. Have you felt worried because of stiffness?	__	__	__	__	
31. Have you felt embarrassed because of stiffness?	__	__	__	__	
32. Has it taken you longer to do your daily activities because of stiffness?	__	__	__	__	
33. Have you been unable to do your daily activities because of stiffness?	__	__	__	__	
34. Have you had to change your plans or behaviour because of stiffness?	__	__	__	__	
35. Have you had to work around stiffness more than usual?	__	__	__	__	
36. Have you had to accept help from others because of stiffness?	__	__	__	__	
37. Have you had to spend more time than usual coping (managing, dealing with, making do) with stiffness?	__	__	__	__	
38. Have you been able to cope (manage, deal, make do) with stiffness?	__	__	__	__	

Some people with RA have told us that to understand the **impact** of a symptom we should ask about its **severity**, its **effect** on the patient's daily life and the patient's ability to **cope (manage, deal with, make do)** with it.

We would like to know the impact that **RA stiffness** has had on your life during the **past 7 days**.

39. Please circle the number that best describes the **severity** of your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

No stiffness 0 1 2 3 4 5 6 7 8 9 10 Extreme stiffness

40. Please circle the number that best describes how well you have been able to **deal** with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

Not well at all 0 1 2 3 4 5 6 7 8 9 10 Very well

41. Please circle the number that best describes how well you have **coped** with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

Not well at all 0 1 2 3 4 5 6 7 8 9 10 Very well

42. Please circle the number that best describes how well you have **managed** your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

Not well at all 0 1 2 3 4 5 6 7 8 9 10 Very well

43. Please circle the number that best describes how well you have **coped (managed, dealt with, made do)** with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

Not well at all 0 1 2 3 4 5 6 7 8 9 10 Very well

44. Please circle the number that best describes the **effect** RA stiffness (in your joints or your body, and at any time of day) has had on your life during the past 7 days

No effect 0 1 2 3 4 5 6 7 8 9 10 Great effect

45. Please circle the number that best describes the overall **impact** on your life of RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

No impact 0 1 2 3 4 5 6 7 8 9 10 Great impact

Some people with RA have told us that to understand the **impact** of a symptom we should ask about its **severity**, its **effect** on the patient's daily life and the patient's ability to **cope (manage, deal with, make do)** with it. We would like to know the impact that **RA stiffness** has had on your life during the **past 7 days**.

46. Please mark the line in the position that best describes the **severity** of your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

No stiffness ——— Extreme stiffness

47. Please circle the number that best describes how well you have been able to **deal** with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

Not well at all ——— Very well

48. Please mark the line in the position that best describes how well you have **coped** with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

Not well at all ——— Very well

49. Please mark the line in the position that best describes how well you have **managed** your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

Not well at all ——— Very well

50. Please mark the line in the position that best describes how well you have **coped (managed, dealt with, made do)** with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

Not well at all ——— Very well

51. Please mark the line in the position that best describes the **effect** RA stiffness (in your joints or your body, and at any time of day) has had on your life during the past 7 days

No effect ——— Great effect

52. Please mark the line in the position that best describes the overall **impact** on your life of RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

No impact ——— Great impact

Some people with RA have told us that to understand the **impact** of a symptom we should ask about its **severity**, its **effect** on the patient's daily life and the patient's ability to **cope (manage, deal with, make do)** with it.

We would like to know the impact that **RA stiffness** has had on your life during the **past 7 days**.

53. Please tick the box that best describes the **severity** of your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

- No stiffness ☐
- A little stiffness ☐
- Quite a bit of stiffness ☐
- Extreme stiffness ☐

54. Please tick the box that best describes how well you have been able to **deal** with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

- Not at all well ☐
- Not very well ☐
- Quite well ☐
- Very well ☐

55. Please tick the box that best describes how well you have been able to **cope** with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

- Not at all well ☐
- Not very well ☐
- Quite well ☐
- Very well ☐

56. Please tick the box that best describes how well you have been able to **manage** your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

- Not at all well ☐
- Not very well ☐
- Quite well ☐
- Very well ☐

57. Please tick the box that best describes how well you have been able to **cope (manage, deal with, make do)** with your RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

- Not at all well ☐
- Not very well ☐
- Quite well ☐
- Very well ☐

58. Please tick the box that best describes the **effect** RA stiffness (in your joints or your body, and at any time of day) has had on your life during the past 7 days

- No effect ☐
- A little effect ☐
- Quite a bit of an effect ☐
- A great effect ☐

59. Please tick the box that best describes the overall **impact** on your life of RA stiffness (in your joints or your body, and at any time of day) during the past 7 days

- No impact ☐
- A little impact ☐
- Quite a bit of an impact ☐
- A great impact ☐

For each of the following questions, please tick <u>one</u> answer that best applies to you

Please indicate how much each of the following have **<u>contributed to your experience</u>** of RA stiffness during the **<u>last 7 days</u>**

	No contri-bution	**Some contri-bution**	**Moderate contri-bution**	**The main contri-bution**
60. The weather	__	__	__	__
61. Being immobile for a period of time for example, in a chair or in bed	__	__	__	__
62. Inflammation or swelling	__	__	__	__
63. Being in an RA flare (flare-up)	__	__	__	__
64. Joint damage	__	__	__	__
65. Other RA symptoms	__	__	__	__
66. Overdoing things for example, doing too much	__	__	__	__
67. RA medications not controlling your disease	__	__	__	__
68. Please specify anything that applies to you but is not listed above ……	__	__	__	__

69. How would you describe the overall level of morning stiffness you have had from the time you wake up?

No stiffness ⊢——————————————⊣ Very severe stiffness

70. How would you describe the overall level of morning stiffness you have had from the time you wake up?

No stiffness 0 1 2 3 4 5 6 7 8 9 10 Very severe stiffness

71. How would you describe the overall level of morning stiffness you have had from the time you wake up?

- ☐ No stiffness
- ☐ Mild stiffness
- ☐ Moderate stiffness
- ☐ Severe stiffness
- ☐ Very severe stiffness

72. How long does your morning stiffness last from waking until maximum improvement occurs?

Minutes _____ or hours _____

73. Were your joints stiff when you woke up today? If yes, how long did this stiffness last?

- ☐ Less than 30 minutes
- ☐ 30 minutes–1 hour
- ☐ 1–2 hours
- ☐ 2–4 hours
- ☐ Over 4 hours
- ☐ All day

74. How long does your morning stiffness last from waking until maximum improvement occurs?

- ☐ Up to 1 hour
- ☐ 1-3 hours
- ☐ ≥3 hours

75. Thinking about a **usual week** when you are **not in a flare** (flare-up) of your RA please circle the number that shows your **usual** RA stiffness (in your joints or your body, and at any time of day)

No stiffness 0 1 2 3 4 5 6 7 8 9 10 Extreme stiffness

76. Thinking about a **usual week** when you are **not in a flare** (flare-up) of your RA please mark on the line your **usual** RA stiffness (in your joints or your body, and at any time of day)

No stiffness ——————————————— Extreme stiffness

77. Thinking about a **usual week** when you are **not in a flare** (flare-up) of your RA please tick the box that best describes your **usual** RA stiffness (in your joints or your body, and at any time of day)

- No stiffness ☐
- Some stiffness ☐
- Quite a lot of stiffness ☐
- Extreme stiffness ☐

Thank you very much for completing this questionnaire!

Appendix T: Probes for cognitive interviews (Study 3)

No.	Probes
Intro.	• What does **active disease** mean to you? • What does **good/bad disease** mean to you? • What does **joint damage** mean to you? • How much would you say you know about **joint damage**? • Are you able to exclude that information? • What period of time is it asking about/what does that cover for you? • How easy/difficult is it to remember stiffness over that period? • Can you tell me in your own words what the intro is telling you?
1	• What does **stiffness in joints** mean to you? • How well does this apply to you? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • Tell me why you chose x rather than y?
2	• What does **stiffness in your body** mean to you? • How well does this apply to you? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • Tell me why you chose x rather than y? • Is it useful or confusing having Q1 and Q2 as different questions?
3	• What does **coming and going as frequently as usual** mean to u? • How well does this apply to you? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • Tell me why you chose x rather than y?
4	• What does **period of immobility** mean to you? • What is the instruction telling you? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • Tell me why you chose x rather than y?
5	• What does **RA medications** mean to you? • How well does this apply to you? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • Tell me why you chose x rather than y?
Inst.	• What is the instruction telling you? • Is that easy or difficult to follow?
6	• What does in your joints and body mean to you? • Tell me what you were thinking when I asked when have you experienced stiffness? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • How well does that question apply to you?
7	• What does in your joints and body mean to you? • Tell me what you were thinking when I asked when have you experienced stiffness? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • How well does that question apply to you? • Is Q6 or Q7 easier to answer?

No.	Probes
Inst.	• What is the instruction telling you? • Is that easy or difficult to follow?
8	• What is the introduction asking you to do? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • Can you tell me in your own words what the question is asking? • How well does this question apply to you?
9	• Can you tell me in your own words what the question is asking?
10	• Can you tell me in your own words what the question is asking?
11	• How well does this apply to you?
12	• What does **daily activities** mean to you?
13	• In your own words what is the question asking?
14	• What does **hobbies or activities you enjoy** mean to you?
15	• In your own words what is the question asking?
16	• What does **rise from a chair** mean to you?
17	• What does **more effort than usual** mean to you?
18	• Would you say it stays the same or varies?
19	• In your own words what is the question asking?
20	• In your own words what is the question asking? • What does **fine movement** mean to you?
21	• In your own words what is the question asking?
22	• How well does this apply to you?
23	• What does **physical strength** mean to you? • How well does this apply to you?
24	• What does **restricted movement** mean to you? • In your own words what is the question asking?
25	• What does **difficult to balance** mean to you? • What does **physically support yourself** mean to you? • How well does this apply to you?
26	• In your own words what is the question asking?
27	• How well does this apply to you? • In your own words what is the question asking?
28	• In your own words what is the question asking?
29	• In your own words what is the question asking?
30	• In your own words what is the question asking?
31	• What does the word **embarrassed** mean to you? • Are their better words e.g. **self-conscious**? • How well does this apply to you?
32	• In your own words what is the question asking?
33	• In your own words what is the question asking?
34	• In your own words what is the question asking?
35	• What does work around mean to you?
36	• How well does this apply to you?
37	• In your own words what is the question asking? • How well does this apply to you?
38	• How well does this apply to you?
Inst.	• What does the term **impact** mean to you? • What does the term **severity** mean to you? • What does the term **effect** mean to you?

No.	Probes
	• What do the terms **cope (manage, deal with, make do)** mean to you? • What is the instruction telling you?
39-45	• What are the instructions telling you to do? • What does **in your joints or your body, and at any time of day** mean to you? • What does the term **severity/deal/cope/manage/effect/impact** mean to you? • How easy or difficult is it to remember the **severity/how you have dealt/coped/managed/effect/impact** of your stiffness over the past 7 days? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • What are your thought about the format of the question?
46-52	• Same as above • What are your thought about the format of the question?
53-59	• Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • What are your thought about the format of the question?
Inst.	• In your own words what is the question asking • What does the phrase **contributed to your experience of RA stiffness** mean to you? • What is the introduction asking you to do?
60-68	• Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • Can you tell me in your own words what the question is asking? • What does the word contribution mean to you? • How well does this question apply to you?
69-71	• What does **not in a flare (flare-up)** mean to you? • What does **usual week** mean to you? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • What are your thought about the format of the question?
72-74	• What does the phrase **overall level of morning stiffness** mean to you? • What does **morning stiffness** mean to you? • In your own words what is the question asking? • How well does the question apply to you? • How easy/difficult is it to remember? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • What are your thought about the format of the question?
75	• In your own words what is the question asking? • What does from **waking to maximal improvement** mean to you? • How well does the question apply to you? • How easy/difficult is it to remember? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • What are your thought about the format of the question?

No.	Probes
76	• In your own words what is the question asking? • How well does the question apply to you? • How easy/difficult is it to remember? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • What are your thought about the format of the question?
77	• In your own words what is the question asking? • How well does the question apply to you? • How easy/difficult is it to remember? • Was it easy or difficult to decide what answer to give? • Do the response options apply to your experience? • What are your thought about the format of the question?
Specific	**Impact triad questions** **Self-management** • Can you tell me if these questions are the same or different for you? • Do they mean the same thing to you? • Do you need them all or just one or some? • What does each word mean to you? **Severity** • What does severity mean to you? Is there a better word (amount/how much/level/other?) **Effect** • What does effect mean to you? Is there a better word (importance/other?) **Other** • Was there anything you thought was irrelevant? • Is it confusing having the instruction change e.g. tick one/tick all that applies? • Would you prefer RA stiffness in ever stem question? (What does it include/exclude?)
Layout	• Do you like the layout? Would you prefer any sections to be in a different order? • Do you like the font? • Are the words big enough? • Are the underlined parts in correct/useful places? • Do you like the small boxes for response options? • Some of the questions had different formats-do you have any preferences for numbers, line or words? • Should the short scales be separate from the main questionnaire body?
General	• In your own words, what is this question asking? • How did you come up with that answer? • Tell me more about that… • What does the term...mean to you? • What time period are you thinking of? • You hesitated a bit there, what are you wondering about? • How could we phrase that question better?
Intro.=introduction; Inst.=instruction	

Appendix U: COREQ checklist (Study 3)

No. Item	Guide questions/description	How was this component addressed
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Interviewer	Which author/s conducted the interview or focus group?	Identified in Section 7.3.5
2. Credentials	What were the researcher's credentials?	The characteristics of the supervisory team were identified in Section 3.4.2 and the researchers (Halls) background was highlighted in Section 3.2.1
3. Occupation	What was their occupation at the time of the study?	
4. Gender	Was the researcher male or female?	
5. Experience and training	What experience or training did the researcher have?	
Relationship with participants		
6. Relationship established	Was a relationship established prior to study commencement?	
7. Participant knowledge of the interviewer	What did the participants know about the researcher?	Identified in Section 7.3.5
8. Interviewer characteristics	What characteristics were reported about the interviewer?	The researchers (Halls) background, experience and research interests were highlighted in Section 3.2.1
Domain 2: study design		
Theoretical framework		
9. Methodological orientation and theory	What methodological orientation was stated to underpin the study?	The methodological approach to the study was discussed in Section 3.3.1 and the method and analysis approach were discussed in Section 7.3
Participant selection		
10. Sampling	How were participants selected?	
11. Method of approach	How were participants approached?	
12. Sample size	How many participants were in the study?	Identified in Sections 7.3.3 and 7.5.1
13. Non-participation	How many people refused to participate or dropped out? Reasons?	
Setting		
14. Setting of data collection	Where was the data collected?	Identified in Section 7.3.5

15. Presence of non-participants	Was anyone else present besides the participants and researchers?	NA
16. Description of sample	What are the important characteristics of the sample?	Identified in Section 7.5.1
Data collection		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Identified in Section 7.3.5 and Appendices S and T
18. Repeat interviews	Were repeat interviews carried out? If yes, how many?	This was not performed
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	Identified in Section 7.3.5
20. Field notes	Were field notes made during and/or after the interview or focus group?	NA
21. Duration	What was the duration of the interviews or focus group?	Identified in Section 7.3.5 and Box 7.1
22. Data saturation	Was data saturation discussed?	Sampling discussed in Section 7.3.3.1
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	This was not performed
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	
25. Description of the coding tree	Did authors provide a description of the coding tree?	Identified in Section 7.4
26. Derivation of themes	Were themes identified in advance or derived from the data?	Identified in Section 7.4
27. Software	What software, if applicable, was used to manage the data?	Identified in Section 7.4
28. Participant checking	Did participants provide feedback on the findings?	This was not performed
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes? Was each quotation identified?	
30. Data and findings consistent	Was there consistency between the data presented and the findings?	Identified in Section 7.5
31. Clarity of major themes	Were major themes clearly presented in the findings?	
32. Clarity of minor themes	Is there a description of diverse cases or minor themes?	

Appendix V: Questionnaire pack (Study 4)

Invitation letter from lead consultant (Study 4)

[INSERT SITE-SPECIFIC HEADER]

Insert date

Dear Sir/Madam,

Developing and testing a rheumatoid arthritis stiffness questionnaire

I am writing to invite you to join a research study which will help us develop and test a new questionnaire to measure stiffness in people with rheumatoid arthritis. I am writing to many patients with rheumatoid arthritis from this department.

Enclosed with this letter is a Patient Information Sheet that explains about the research study. We would be grateful if you could read it. Taking part in research is quite voluntary. If you would like to take part, please complete the questionnaire and return it in the envelope provided. This will go directly to Serena Halls, the researcher running the study, and I will not know who decides to take part.

Thank you for considering taking part in this study.

Yours sincerely,

Insert consultant name
Insert consultant title/role

Patient information sheet (Study 4)

Developing and testing a rheumatoid arthritis stiffness questionnaire

Patient information sheet

You are being invited to take part in a research study. Before you decide whether to take part, it is important for you to understand why the research is being done, and what it will involve. Please take time to read the following information leaflet carefully and discuss it with others if you wish. If anything is not clear or you would like more information please ask one of the study team.

What is the purpose of the study?
Stiffness is a common problem for people with rheumatoid arthritis (RA). Until recently, stiffness was not very well understood but with the help of people like you we have developed a better understanding of what stiffness is and what it means to people. We would now like your help to develop this further by working towards a better way of measuring stiffness. Stiffness measurement currently is not very accurate and has not been developed with patient input. The aim of this study is to create and then test a new RA stiffness questionnaire that has been developed based on the patient experience of stiffness.

Why have I been chosen?
You have been chosen because you have a diagnosis of RA. We are inviting people with RA from Rheumatology Departments in the UK to take part. We would like a range of people to complete the questionnaire including those who have lots of stiffness, those who have a little bit of stiffness and even those who have none at all. This is because the new questionnaire will need to be able to detect different levels of stiffness.

What will I be asked to do if I take part?
The study involves completing a questionnaire pack. The questionnaire pack includes the stiffness questionnaire that we would like to test and some other questions about you and your arthritis in general. Once you have completed the questionnaire pack, please return it to us in the prepaid envelope, or if you are completing it in clinic, please hand the envelope to the receptionist. We will send a written reminder to people who have not returned the questionnaire pack within 3 weeks.

At the end of the questionnaire pack we have asked whether you would like to receive a summary of the study findings. For this, we will ask you to provide your name and address. This will be kept separately from your returned questionnaire.

We would also like to explore the relationship between the answers people give to the stiffness questionnaire and their disease activity. To do this we would like permission to access your relevant medical records to record the results of your most recent blood test. We will not ask you to undergo a blood test to participate in this study.

Who is asking me to take part?
I am Serena Halls, a PhD research student at the University of the West of England (UWE). This research study is the last of three studies which form my PhD. I am based at the Academic Rheumatology Department in the Bristol Royal Infirmary.

Do I have to take part?
No, you do not have to take part. If you decide not to take part you do not need to give a reason, nobody will be upset and the standard of care you receive will not be affected. Even if you do decide to take part, you are still free to withdraw from the study at any time and do not need to give a reason for doing so.

What are the risks or benefits of taking part in the study?
We do not believe there are any risks in being involved in this study. Although there are no direct benefits to you in taking part, you will be helping us to gain a better understanding of how to measure RA stiffness. This could help us to improve decisions made about treatment and management of RA.

For general advice about taking part in research, or if you have any concerns or complaints about the conduct of the study, you can contact the local Patient Advice and Liaison Service (PALS) by phone: 0117 900 3433, email: pals@bristolpct.nhs.uk or post: PALS, NHS Bristol, South Plaza, Marlborough Street, Bristol, BS1 3NX.

Will my taking part in the study be kept confidential?
Yes. Your name will be replaced by a code by the researcher (Serena Halls). No one will be able to identify you from the questionnaire or study report. The study records will be kept securely for 6 years and then destroyed, in accordance with best practice in research guidelines.

What will happen to the results of this research study?
We hope to report the results at conferences and in professional journals. We will also send you a summary of the results at the end of the study if you would like.

Who is organising, funding and reviewing the research?
The study is coordinated by a team from UWE based at the Academic Rheumatology Unit at the Bristol Royal Infirmary. It is funded by UWE. This study has been reviewed and approved by Wales Research Ethics Committee 4 and the University Research Ethics Committee.

What do I do now?
Thank you for considering taking part in this study. If you would like to take part, please complete the enclosed questionnaire and return it to us in the prepaid, stamped addressed envelope.

Research team:
Serena Halls, PhD Researcher, UWE Bristol
Professor Sarah Hewlett, Professor of Rheumatology Nursing, UWE Bristol
Professor John Kirwan, Emeritus Professor of Rheumatic Diseases, UoB Bristol
Dr Jon Pollock, Reader in Epidemiology, UWE Bristol
Dr Emma Dures, Senior Research Fellow, UWE Bristol
Mrs Avis Edmunds, Patient Research Partner
Mrs Gill Baker, Patient Research Partner

Serena Halls
Academic Rheumatology Unit,
Bristol Royal Infirmary,
Bristol, BS2 8HW
0117 342 4972
Serena.halls@uwe.ac.uk

Professor Sarah Hewlett
Academic Rheumatology Unit,
Bristol Royal Infirmary,
Bristol, BS2 8HW
0117 342 2903
Sarah.hewlett@uwe.ac.uk

Questionnaire booklet (Study 4)

Developing and testing a rheumatoid arthritis stiffness questionnaire

Introduction

We would like to know how rheumatoid arthritis (RA) stiffness has affected you in the past 7 days. We would like you to fill this questionnaire out whether you have had lots of stiffness, a little bit of stiffness or even no stiffness at all.

This questionnaire pack will take about 10 – 20 minutes to complete. When you are ready to begin, please turn overleaf and complete the consent form on page 2. If you would like to keep a copy, please also fill in the second consent form on page 3 which you can tear or cut out for your records.

After you have completed the consent form, please continue through the questionnaire pack and answer all of the questions. You may notice that some of the questions are very similar. This is because we need to test different versions of some of the questions to see which works best, so please answer them all. Don't think too long and hard, just give your first reaction - there are no right or wrong answers!

Thank you

Today's date: _____ Time started: _____ (Hour) _____ (Minute)

Official use only

Study number: _____

[INSERT SITE-SPECIFIC HEADER]

Developing and testing a rheumatoid arthritis stiffness questionnaire

Consent form

Please initial each box

1. I confirm that I have read and understand the information sheet dated 10/12/14 (version 1.2) for the above study. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without my medical care or legal rights being affected ☐
3. I understand information from the questionnaire will be anonymised and may be used in publications, conference presentations and in a PhD thesis ☐
4. I understand that relevant sections of my medical notes and study data may be accessed (in confidence) by the study research team, regulatory authorities or relevant members of the NHS Trust. I give permission for these individuals to have access to my records ☐
5. I agree to take part in the above study ☐

When you have initialled all of the boxes above, please complete the following two lines yourself, including the date.

Name (please print).....

Signature.....Date.....

Name or researcher taking consent.....

Signature .....Date.....

1 copy for patient, 1 for researcher, 1 for hospital notes

[INSERT SITE-SPECIFIC HEADER]

Developing and testing a rheumatoid arthritis stiffness questionnaire

Consent form

Please initial each box

1. I confirm that I have read and understand the information sheet dated 10/12/14 (version 1.2) for the above study. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without my medical care or legal rights being affected ☐
3. I understand information from the questionnaire will be anonymised and may be used in publications, conference presentations and in a PhD thesis ☐
4. I understand that relevant sections of my medical notes and study data may be accessed (in confidence) by the study research team, regulatory authorities or relevant members of the NHS Trust. I give permission for these individuals to have access to my records ☐
5. I agree to take part in the above study ☐

When you have initialled all of the boxes above, please complete the following two lines yourself, including the date.

Name (please print).....

Signature.....Date.....

Name or researcher taking consent.....

Signature .....Date.....

1 copy for patient, 1 for researcher, 1 for hospital notes

Please tell us about your RA

Overall

1. Considering all the ways that your arthritis affects you, mark an X on the scale for how well you are doing

Very well ├──┤ Very poor

Level of pain

2. Please circle the number which shows how much pain you have had in the past 7 days.

No pain	0	1	2	3	4	5	6	7	8	9	10	Worst possible pain

Level of fatigue

3. Please circle the number which shows your average level of fatigue during the past 7 days.

No fatigue	0	1	2	3	4	5	6	7	8	9	10	Totally exhausted

Flare

4. Are you having a flare (flare-up) of rheumatoid arthritis at this time? (Please tick)

No ☐ Yes ☐

Stiffness

5. Please circle the number that best describes the severity of your RA stiffness over a usual week when you are not in a flare?

No stiffness	0	1	2	3	4	5	6	7	8	9	10	Extreme stiffness

Please tell us about your abilities this week

Please tick the one response which best describes your usual ABILITIES over the PAST WEEK

	Without ANY difficulty	With SOME difficulty	With MUCH difficulty	Unable to do
1 DRESSING AND GROOMING Are you able to: - Dress yourself, including tying shoelaces and doing buttons?	_____	_____	_____	_____
2 RISING Are you able to: - Get in and out of bed?	_____	_____	_____	_____
3 EATING Are you able to: - Lift a full cup or glass to your mouth?	_____	_____	_____	_____
4 WALKING Are you able to: - Walk outdoors on flat ground?	_____	_____	_____	_____
5 HYGIENE Are you able to: - Wash and dry your entire body?	_____	_____	_____	_____
6 REACH Are you able to: - Bend down to pick up clothing from the floor?	_____	_____	_____	_____
7 GRIP Are you able to: - Turn taps on and off?	_____	_____	_____	_____
8 ACTIVITIES Are you able to: - Get in and out of a car?	_____	_____	_____	_____

Please tell us how active your arthritis is today

1. In general, how active is your arthritis today?
Mark X on the scale below at the point that best describes the level of arthritis activity

Arthritis inactive ——————————————— Arthritis very active

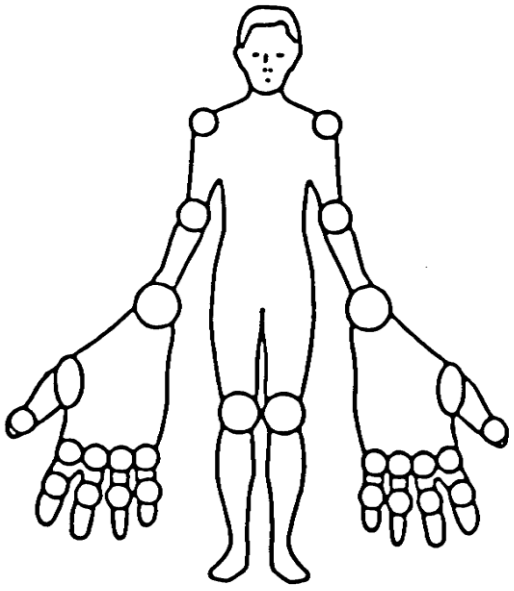
2. Were your joints stiff when you woke up today? No ☐ Yes ☐

If yes, how long did this extra stiffness last?

- Less than 30 minutes ☐
- 30 minutes–1 hour ☐
- 1–2 hours ☐
- 2–4 hours ☐
- Over 4 hours ☐
- All day ☐

3. Swollen joints

Please indicate with a mark, on the picture below all the joints that are <u>SWOLLEN</u> at the present:



Please tell us about your RA stiffness

This questionnaire is about RA stiffness that comes and goes. It is not about joints that are permanently stuck (for example, due to an operation). However, we do appreciate that sometimes even permanently stuck joints do get stiffer (for example, when your disease is bad). Please just try think about the stiffness that comes and goes as you answer this questionnaire.

1. Do you have any joints that are permanently stuck?
   - No ☐
   - Yes ☐

We would like to know how RA stiffness has affected you over the past 7 days

2. Over the past 7 days when have you experienced **RA stiffness**? Please tick all that apply to you
   - In the night ☐
   - In the morning ☐
   - In the afternoon ☐
   - In the evening ☐
   - None of these ☐

For each of the following questions, please tick the one answer that best applies to you

3. Have you experienced **RA stiffness** in your joints over the past 7 days?
   - No, not in any of my joints ☐
   - Yes, in a few of my joints ☐
   - Yes, in many of my joints ☐
   - Yes, in all of my joints ☐

4. Over the past 7 days have you experienced **RA stiffness** all over?
   - No ☐
   - Yes ☐

5. Over the past 7 days has your **RA stiffness** been different to usual for you?

   - It has been much better than usual ☐
   - It has been better than usual ☐
   - It has been the same as usual ☐
   - It has been worse than usual ☐
   - It has been much worse than usual ☐

6. Over the past 7 days has your **RA stiffness** been as variable (coming and going) as usual for you?

   - It has been much less variable than usual ☐
   - It has been less variable than usual ☐
   - It has been the same as usual ☐
   - It has been more variable than usual ☐
   - It has been much more variable than usual ☐

7. Over the past 7 days have you experienced **RA stiffness** after a period of immobility (for example, after sitting for a while)?

   - No, not at all ☐
   - Yes, a little ☐
   - Yes, quite a lot ☐
   - Yes, very much ☐

8. Have you experienced **RA stiffness** in your body (outside of your joints) over the past 7 days?

   - No, not in any part of my body ☐
   - Yes, in a few parts of my body ☐
   - Yes, in many parts of my body ☐
   - Yes, all over my body ☐

For each of the following questions, please tick the one answer that best applies to you

Thinking about your **RA stiffness** over the past 7 days

	Not at all	A little	Quite a lot	Very much
9. Has **RA stiffness** affected your sleep?				
10. Has **RA stiffness** made it difficult to dress or undress yourself?				
11. Has **RA stiffness** made it difficult to wash yourself (for example, have a shower)?				
12. Has **RA stiffness** made it difficult to carry out your responsibilities or commitments?				
13. Has **RA stiffness** made it difficult to do your daily tasks or activities?				
14. Has **RA stiffness** made it difficult to chew?				
15. Has **RA stiffness** made it difficult to do hobbies or activities you enjoy?				
16. Has **RA stiffness** made it difficult to get out of bed?				
17. Has **RA stiffness** made it difficult to get up after sitting for a while?				
18. Have your daily tasks and activities required more effort because of **RA stiffness**?				
19. Has **RA stiffness** made you slower (for example, unable to do things quickly)?				
20. Has **RA stiffness** made it difficult to do fine movements (for example, write with a pen)?				
21. Has **RA stiffness** made it difficult to grip or hold things?				

Thinking about your **RA stiffness** over the past 7 days

	Not at all	**A little**	**Quite a lot**	**Very much**
22. Has **RA stiffness** made it difficult to open and close your fist?				
23. Has **RA stiffness** reduced your strength to do tasks?				
24. Has your movement been restricted because of **RA stiffness**?				
25. Has **RA stiffness** made it difficult to balance without physically supporting yourself?				
26. Have you had to concentrate to move your body because of **RA stiffness**?				
27. Have you felt frustrated because of **RA stiffness**?				
28. Have you felt worried or concerned because of **RA stiffness**?				
29. Have you felt self-conscious because of **RA stiffness**?				
30. Has it taken you longer to do your daily tasks or activities because of **RA stiffness**?				
31. Have you had to change your plans or behaviour because of **RA stiffness**?				
32. Have you had to work around your **RA stiffness** (or do things in a different way)?				
33. Have you needed help (from others or gadgets) because of **RA stiffness**?				

34. Please circle the number that best describes the impact that **RA stiffness** has had on your life over the past 7 days

No impact at all	0	1	2	3	4	5	6	7	8	9	10	A great deal of impact

35. Please circle the number that best describes the severity of your **RA stiffness** over the past 7 days

No stiffness	0	1	2	3	4	5	6	7	8	9	10	Extreme stiffness

36. Please circle the number that best describes how important **RA stiffness** has been in your life over the past 7 days

Not important at all	0	1	2	3	4	5	6	7	8	9	10	Very important

37. Please circle the number that best describes how well you have coped with your **RA stiffness** over the past 7 days

Not well at all	0	1	2	3	4	5	6	7	8	9	10	Very well

38. How much of the stiffness you have reported in the questions above is about joints that are permanently stuck?

- None of the stiffness I have reported ☐
- A little of the stiffness I have reported ☐
- Quite a lot of the stiffness I have reported ☐
- All of the stiffness I have reported ☐

These questions are also about stiffness. You may be familiar with them from clinic or from other questionnaires. We would now like to formally test them in this study.

1. How would you describe the overall level of morning stiffness you have had from the time you wake up?

No stiffness	0	1	2	3	4	5	6	7	8	9	10	Very severe stiffness

2. How long does your morning stiffness last from waking until maximum improvement occurs?

   Up to 1 hour ☐

   1 – 3 hours ☐

   More than 3 hours ☐

3. Circle the number that best describes the stiffness (all over or in your joints) you felt due to your rheumatoid arthritis during the last week:

No stiffness	0	1	2	3	4	5	6	7	8	9	10	Extreme stiffness

4. How would you describe the overall level of morning stiffness you have had from the time you wake up?

   No stiffness ☐

   Mild stiffness ☐

   Moderate stiffness ☐

   Severe stiffness ☐

   Very severe stiffness ☐

5. How long does your morning stiffness last from waking until maximum improvement occurs?

   Minutes _____ or hours _____

Please tell us about you

1. Are you male or female? (Please tick) Male ☐ Female ☐

2. What is your date of birth? Day ☐ Month ☐ Year ☐

3. Approximately how long have you had rheumatoid arthritis (RA)? Years ☐

4. What medications are you taking for your RA? (Please write below)

_____

_____

_____

5. Do you have any other medical conditions for which you are receiving treatment? (Please write below)

_____

_____

_____

6. What is your work status (Please tick)

- ☐ Paid work
- ☐ Student
- ☐ Homemaker
- ☐ Unemployed
- ☐ Retired
- ☐ Receiving incapacity benefits

7. What is your level of education (Please tick highest level)

- ☐ Did not complete school
- ☐ School education
- ☐ College / apprenticeship
- ☐ University level education

8. What is your postcode ☐

If you would like to receive a summary of the results of this study, please tick the box below and provide your name and address.

The information you provide here will be kept confidential and will be stored separately from your returned questionnaire in a locked cabinet. You will not be contacted for any other reason.

Yes, I would like to receive a summary of the results ☐

Your name: _____

Your address: _____

_____

_____

Thank you very much for taking the time to fill in this questionnaire!

Please now place it in the prepaid envelope provided and post back to us

Research team:
Serena Halls, PhD Researcher, UWE Bristol
Professor Sarah Hewlett, Professor of Rheumatology Nursing, UWE Bristol
Professor John Kirwan, Emeritus Professor of Rheumatic Diseases, UoB Bristol
Dr Jon Pollock, Reader in Epidemiology, UWE Bristol
Dr Emma Dures, Senior Research Fellow, UWE Bristol
Mrs Avis Edmunds, Patient Research Partner
Mrs Gill Baker, Patient Research Partner

Serena Halls
Academic Rheumatology Unit,
Bristol Royal Infirmary,
Bristol, BS2 8HW
0117 342 4972
Serena.halls@uwe.ac.uk

Professor Sarah Hewlett
Academic Rheumatology Unit,
Bristol Royal Infirmary,
Bristol, BS2 8HW
0117 342 2903
Sarah.hewlett@uwe.ac.uk

Appendix W: Frequency and distribution graphs (normality assessment where relevant) for non-stiffness items (Study 4)

Whole sample: Age

The variable age for the whole sample was treated as a continuous variable and the extent to which it met the assumption of normality was explored. Negative skew (-0.407, SE=0.096) and kurtosis (-0.298, SE=0.192) and a significant Kolmogorov-Smirnov statistic (D(643)=0.069, *p*=0.000) indicated violation of the assumption of normality. However, in large samples (>200) these tests are too sensitive and results such as these are common, thus inspection of the histogram is recommended (Tabachnick and Fidell, 2007). The histogram appeared to have a reasonably normal distribution and the values on the Q-Q plot fell near to or on the straight line (Figure W.1).

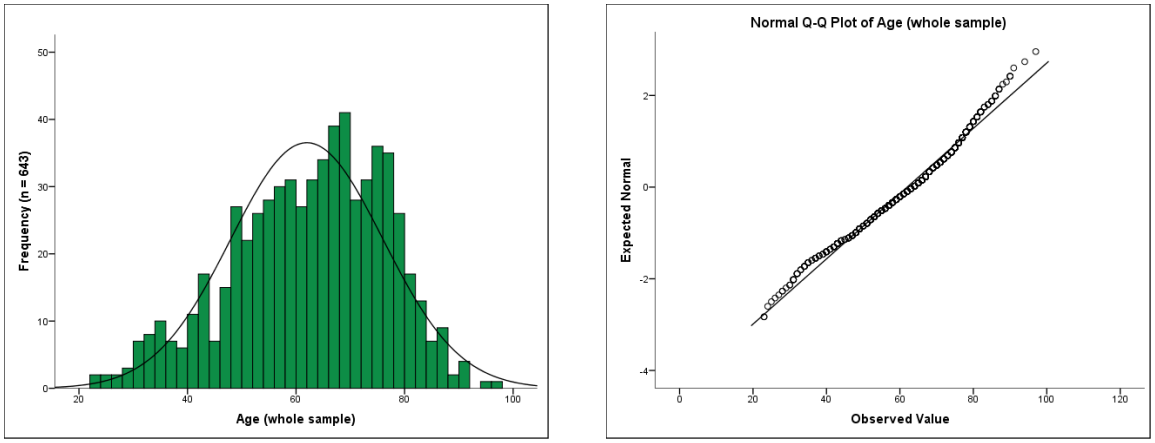


Figure W.1: Histogram and Q-Q plot for age (whole sample)

Non-responders: Age

The variable age for the non-responders was explored for the extent to which it met the assumption of normality. Negative skew (-0.350, SE=0.128) and kurtosis (-0.518, SE=0.254) and a significant Kolmogorov-Smirnov statistic (D(366)=0.073, *p*=0.000) indicated violation of the assumption of normality. The histogram appeared to have a reasonably normal distribution and the values on the Q-Q plot fell near to or on the straight line (Figure W.2).

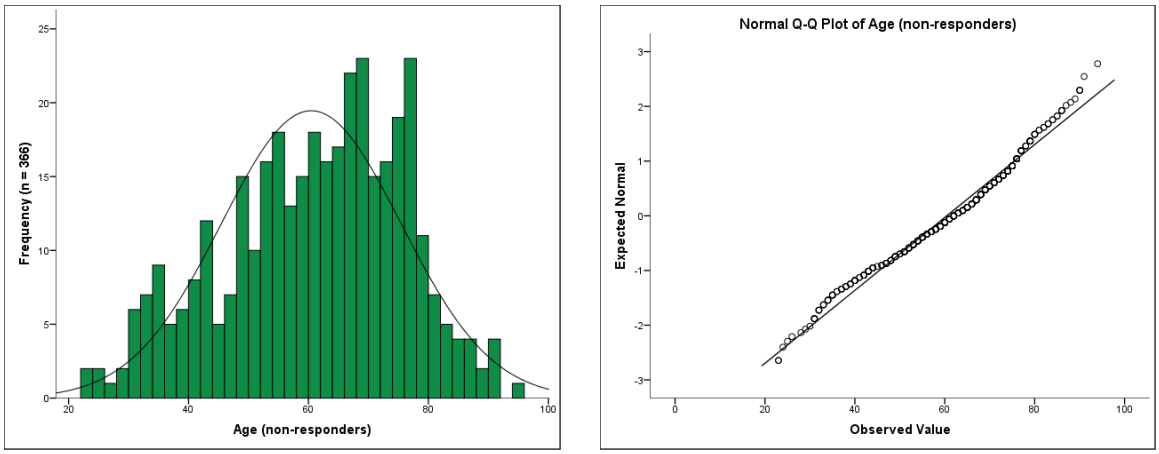


Figure W.2: Histogram and Q-Q plot for age (non-responders)

Responders: Age

The variable age for the responders was explored for the extent to which it met the assumption of normality. Negative skew (-0.330, SE=0.146) and kurtosis (-0.195, SE=0.292) and a significant Kolmogorov-Smirnov statistic (D(277)=0.058, *p*=0.000) indicated violation of the assumption of normality. The histogram appeared to have a reasonably normal distribution and the values on the Q-Q plot fell near to or on the straight line (Figure W.3).

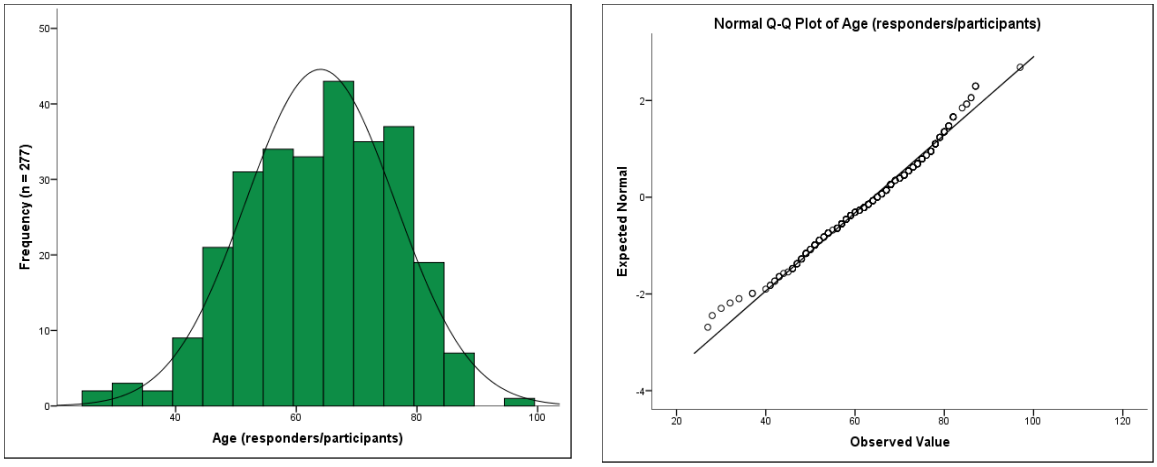


Figure W.3: Histogram and Q-Q plot for age (responders)

Responders: Disease duration

Responder disease duration was explored for the extent to which it met the assumption of normality. Positive skew (1.337, SE=0.148) and kurtosis (1.084, SE=0.295) and a significant Kolmogorov-Smirnov statistic (D(271)=0.203, *p*=0.000) indicated violation of the assumption of normality. The histogram was not normally distributed and the values on the Q-Q plot fell in an s-shape around the straight line (Figure W.4).

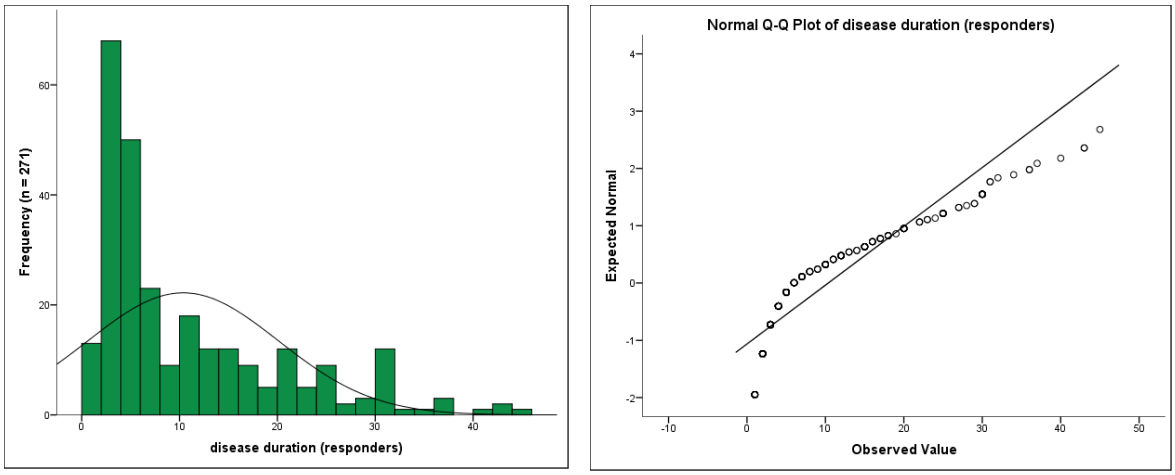


Figure W.4: Histogram and Q-Q plot for disease duration (responders)

Responders: Disease activity (PDAS2, Choy *et al,* 2008)

Responder disease activity (PDAS2, Choy *et al,* 2008) was explored for the extent to which it met the assumption of normality. Positive skew (0.405, SE=0.146) and negative kurtosis (-0.786, SE=0.292) and a significant Kolmogorov-Smirnov statistic (D(277)=0.094, *p*=0.000) indicated violation of the assumption of normality. The

histogram was not normally distributed and the values on the Q-Q plot dropped off the straight line at the ends (Figure W.5).

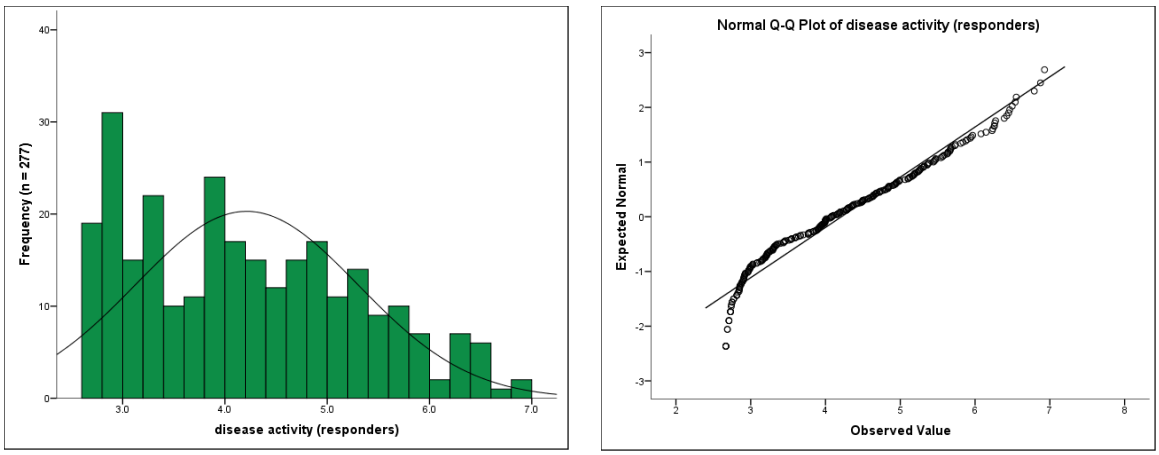


Figure W.5: Histogram and Q-Q plot for disease activity (responders)

Responders: Disability (MHAQ, Pincus *et al,* 1983)

Responder disability scores (MHAQ, Pincus *et al,* 2008) was explored for the extent to which it met the assumption of normality. Positive skew (0.796, SE=0.146) and negative kurtosis (-0.117, SE=0.292) and a significant Kolmogorov-Smirnov statistic (D(277)=0.146, *p*=0.000) indicated violation of the assumption of normality. The histogram was not normally distributed and the values on the Q-Q plot did not fall on the straight line (Figure W.6).

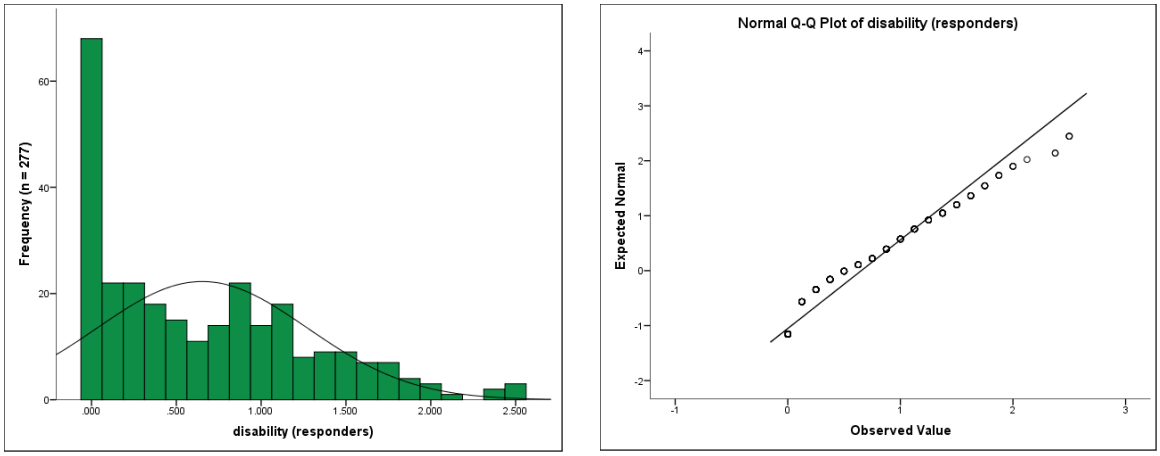


Figure W.6: Histogram and Q-Q plot for disability (responders)

Responders: Patient global assessment (PtG)

Patient global assessment (PtG) was explored for the extent to which it met the assumption of normality. Positive skew (0.337, SE=0.147) and negative kurtosis (-0.892, SE=0.293) and a significant Kolmogorov-Smirnov statistic (D(275)=0.079, *p*=0.000) indicated violation of the assumption of normality. The histogram was not normally distributed and the values on the Q-Q plot dropped off the straight line at the ends (Figure W.7).

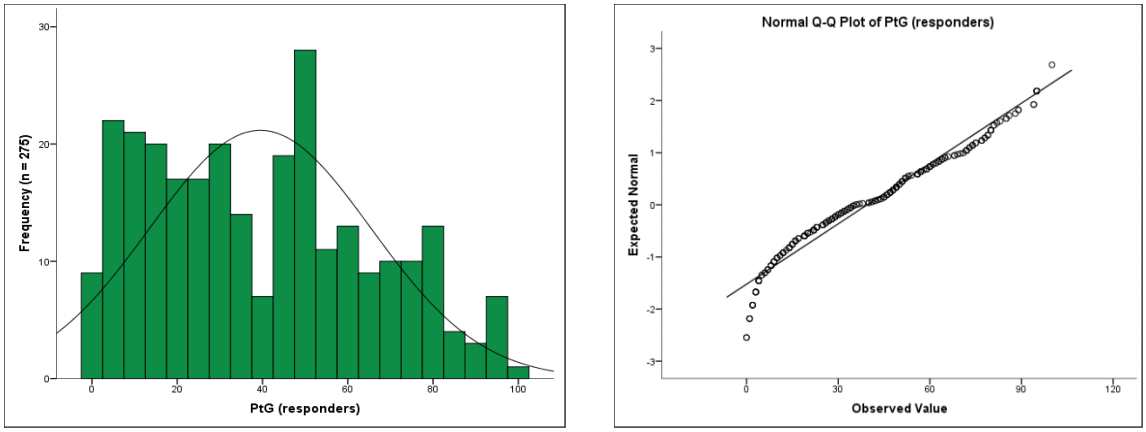


Figure W.7: Histogram and Q-Q plot for PtG (responders)

Responders: Pain

Responder pain (NRS) was explored for the extent to which it met the assumption of normality. Negative skew (-0.025, SE=0.146) and kurtosis (-1.106, SE=0.292) and a significant Kolmogorov-Smirnov statistic (D(277)=0.120, *p*=.000) indicated violation of the assumption of normality. The histogram was quite flat and the values on the Q-Q plot fell around the straight line (Figure W.8).

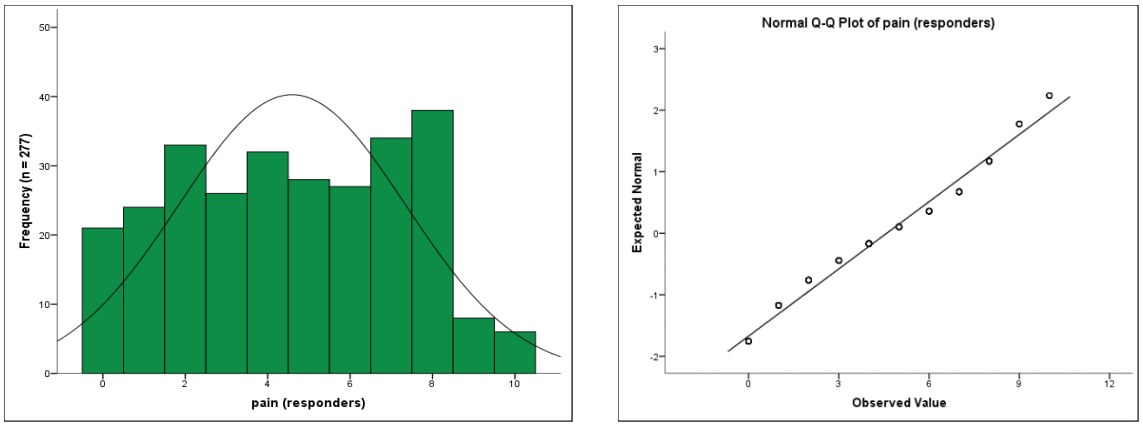


Figure W.8: Histogram and Q-Q plot for pain (responders)

Responders: Fatigue (BRAF severity NRS, Nicklin *et al,* 2010a, Nicklin *et al,* 2010b)

Responder fatigue (BRAF severity NRS, Nicklin *et al,* 2010a, Nicklin *et al,* 2010b) was explored for the extent to which it met the assumption of normality. Negative skew (-0.418, SE=0.146) and kurtosis (-0.842, SE=0.292) and a significant Kolmogorov-Smirnov statistic (D(277)=0.161, *p*=0.000) indicated violation of the assumption of normality. The histogram was quite flat and the values on the Q-Q plot fell around the straight line (Figure W.9).

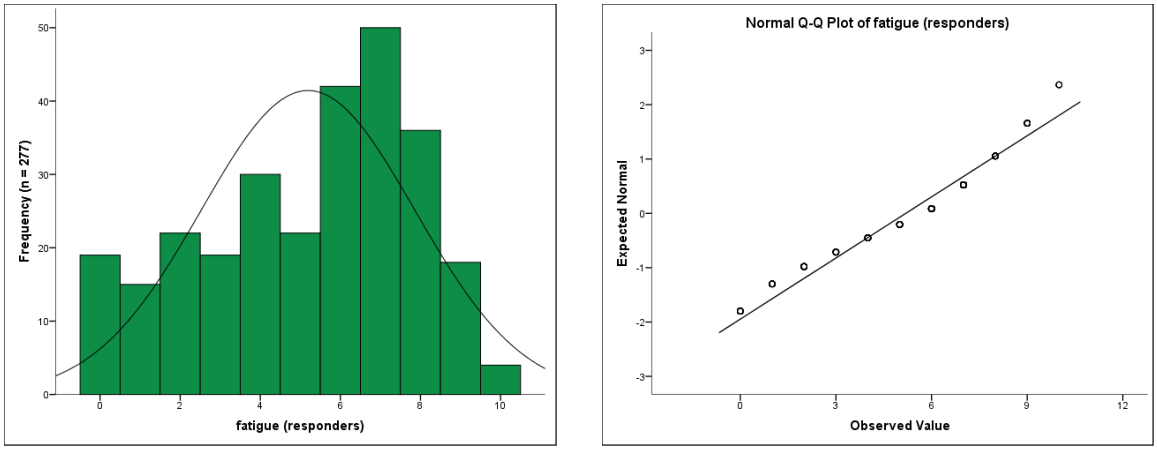


Figure W.9: Histogram and Q-Q plot for fatigue (responders)

Appendix X: Pearson's correlation coefficient (Study 4)

	7.2	7.3	7.4	8.5	8.7	8.8	9.9	9.10	9.11	9.12	9.13	9.15	9.16	9.17	9.18	9.19	9.20	9.21	10.22
7.2	1.000	.579	.388	.373	.578	.479	.536	.491	.437	.487	.541	.464	.443	.514	.544	.482	.474	.471	.435
7.3	-	1.000	.436	.386	.548	.465	.549	.530	.480	.526	.557	.518	.509	.547	.581	.593	.504	.537	.441
7.4	-	-	1.000	.221	.398	.516	.407	.345	.311	.338	.343	.358	.373	.346	.331	.308	.299	.297	.297
8.5	-	-	-	1.000	.368	.315	.345	.377	.321	.398	.457	.361	.290	.294	.399	.397	.371	.331	.328
8.7	-	-	-	-	1.000	.527	.528	.560	.482	.506	.550	.541	.573	.696	.578	.539	.462	.496	.397
8.8	-	-	-	-	-	1.000	.491	.462	.360	.467	.513	.509	.470	.485	.532	.483	.413	.432	.276
9.9	-	-	-	-	-	-	1.000	.602	.586	.586	.627	.619	.625	.568	.659	.623	.592	.602	.450
9.10	-	-	-	-	-	-	-	1.000	.833	.761	.782	.710	.751	.680	.698	.671	.667	.648	.555
9.11	-	-	-	-	-	-	-	-	1.000	.770	.749	.671	.708	.613	.669	.633	.638	.656	.555
9.12	-	-	-	-	-	-	-	-	-	1.000	.856	.761	.661	.625	.787	.736	.629	.721	.631
9.13	-	-	-	-	-	-	-	-	-	-	1.000	.783	.660	.636	.821	.760	.678	.723	.602
9.15	-	-	-	-	-	-	-	-	-	-	-	1.000	.668	.666	.799	.746	.640	.652	.547
9.16	-	-	-	-	-	-	-	-	-	-	-	-	1.000	.719	.667	.619	.588	.607	.456
9.17	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	.695	.654	.556	.567	.444
9.18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	.836	.653	.717	.591
9.19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	.673	.751	.617
9.20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	.756	.641
9.21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	.666
10.22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000
10.23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.26	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.28	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10.33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11.34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11.35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11.36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11.37	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12.4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

All significant at $p<.01$

	10.23	10.24	10.25	10.26	10.27	10.28	10.29	10.30	10.31	10.32	10.33	11.34	11.35	11.36	11.37	6.2	12.1	12.3	12.4
7.2	.457	.510	.313	.396	.480	.383	.472	.472	.386	.456	.435	.601	.620	.548	-.250	.479	.587	.618	.590
7.3	.521	.511	.407	.424	.500	.357	.387	.532	.402	.483	.444	.614	.640	.587	-.251	.518	.628	.630	.632
7.4	.197	.288	.300	.237	.258	.347	.336	.327	.326	.297	.308	.284	.380	.387	-.127	.218	.345	.352	.371
8.5	.335	.392	.297	.310	.317	.310	.264	.364	.309	.304	.287	.497	.541	.466	-.238	.422	.439	.521	.489
8.7	.510	.619	.419	.518	.568	.403	.520	.589	.459	.459	.503	.634	.616	.614	-.243	.459	.632	.634	.581
8.8	.448	.484	.353	.409	.453	.404	.372	.468	.403	.427	.416	.548	.553	.515	-.228	.445	.503	.514	.483
9.9	.551	.588	.473	.519	.610	.499	.499	.619	.563	.583	.563	.671	.638	.643	-.315	.496	.655	.627	.626
9.10	.630	.731	.612	.622	.655	.521	.545	.698	.609	.681	.674	.674	.680	.666	-.288	.523	.715	.670	.676
9.11	.611	.689	.659	.627	.618	.502	.539	.676	.604	.662	.654	.636	.629	.613	-.304	.481	.653	.623	.616
9.12	.673	.716	.621	.594	.664	.573	.575	.767	.732	.744	.647	.739	.693	.699	-.349	.573	.713	.696	.676
9.13	.681	.739	.621	.612	.680	.534	.560	.796	.678	.733	.697	.777	.763	.737	-.350	.556	.739	.745	.700
9.15	.666	.755	.621	.643	.736	.654	.595	.765	.699	.725	.638	.759	.718	.733	-.285	.569	.699	.701	.636
9.16	.588	.639	.607	.665	.599	.535	.531	.611	.552	.612	.591	.653	.625	.650	-.255	.480	.714	.622	.647
9.17	.610	.693	.637	.666	.661	.500	.553	.683	.542	.611	.579	.694	.660	.662	-.311	.500	.708	.682	.643
9.18	.728	.775	.616	.699	.731	.632	.610	.860	.714	.745	.656	.806	.779	.771	-.372	.625	.740	.754	.699
9.19	.763	.761	.631	.667	.716	.621	.617	.844	.721	.752	.714	.796	.769	.787	-.354	.639	.739	.750	.713
9.20	.688	.662	.537	.568	.595	.492	.551	.670	.612	.662	.654	.648	.620	.634	-.322	.491	.597	.617	.576
9.21	.786	.701	.563	.592	.632	.515	.550	.734	.624	.692	.727	.679	.650	.675	-.361	.496	.646	.662	.615
10.22	.688	.638	.457	.503	.583	.460	.468	.600	.509	.568	.521	.602	.566	.554	-.291	.494	.557	.549	.570
10.23	1.000	.761	.590	.635	.688	.593	.622	.760	.645	.751	.738	.711	.670	.685	-.360	.534	.619	.653	.623
10.24	-	1.000	.683	.710	.746	.615	.645	.800	.699	.731	.710	.757	.748	.743	-.333	.565	.698	.749	.667
10.25	-	-	1.000	.755	.591	.456	.524	.657	.569	.588	.636	.603	.601	.562	-.262	.461	.564	.613	.532
10.26	-	-	-	1.000	.654	.562	.595	.697	.612	.636	.605	.660	.651	.648	-.307	.517	.601	.642	.558
10.27	-	-	-	-	1.000	.738	.748	.786	.726	.760	.683	.764	.728	.783	-.355	.568	.682	.720	.645
10.28	-	-	-	-	-	1.000	.726	.644	.655	.653	.552	.645	.600	.672	-.343	.495	.556	.583	.540
10.29	-	-	-	-	-	-	1.000	.677	.676	.677	.659	.644	.593	.658	-.266	.458	.586	.600	.527
10.30	-	-	-	-	-	-	-	1.000	.787	.823	.753	.814	.770	.786	-.336	.610	.718	.762	.683
10.31	-	-	-	-	-	-	-	-	1.000	.830	.687	.704	.645	.705	-.284	.508	.602	.638	.553
10.32	-	-	-	-	-	-	-	-	-	1.000	.714	.720	.673	.704	-.306	.486	.651	.673	.629
10.33	-	-	-	-	-	-	-	-	-	-	1.000	.689	.657	.680	-.291	.459	.633	.662	.577
11.34	-	-	-	-	-	-	-	-	-	-	-	1.000	.920	.921	-.338	.703	.878	.906	.810
11.35	-	-	-	-	-	-	-	-	-	-	-	-	1.000	.908	-.344	.711	.872	.934	.831
11.36	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	-.360	.664	.867	.888	.759
11.37	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	-.247	.670	.677	.664
6.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	-.323	-.337	-.373
12.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	.891	.864
12.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000	.839
12.4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.000

All significant at $p<.01$

Appendix Y: Oblique promax rotation pattern and structure matrices (Study 4)

PCA

Items	Rotated component loadings							
	Pattern matrix				Structure matrix			
	1	2	3	4	1	2	3	4
9.11 draft stiffness item 11	**1.078**	-.133	-.163	.030	**.868**	.603	.594	.517
9.10 draft stiffness item 10	**.944**	-.040	-.164	.148	**.876**	.667	.616	.611
9.21 draft stiffness item 21	**.796**	.143	.024	-.138	**.841**	.686	.670	.451
9.12 draft stiffness item 12	**.744**	.135	.055	-.039	**.867**	.723	.707	.530
10.25 draft stiffness item 25	**.837**	-.190	.118	-.015	**.773**	.530	.625	.435
9.13 draft stiffness item 13	**.712**	.244	-.029	.002	**.878**	.774	.696	.577
9.20 draft stiffness item 20	**.774**	.129	-.019	-.093	**.802**	.653	.623	.451
9.16 draft stiffness item 16	**.788**	-.198	-.066	.362	**.805**	.596	.594	.681
10.33 draft stiffness item 33	**.618**	-.104	.366	-.029	**.803**	.613	.755	.472
10.24 draft stiffness item 24	**.553**	.102	.287	.000	**.854**	.732	.788	.552
10.23 draft stiffness item 23	**.579**	.112	.304	-.137	**.817**	.686	.760	.446
10.26 draft stiffness item 26	**.591**	-.125	.314	.058	**.773**	.590	.714	.502
10.22 draft stiffness item 22	**.653**	.401	-.039	-.362	**.711**	.644	.561	.272
9.18 draft stiffness item 18	**.440**	.237	.265	.038	**.851**	.789	.794	.597
9.15 draft stiffness item 15	**.477**	.022	.328	.121	**.822**	.699	.776	.597
9.19 draft stiffness item 19	**.403**	.308	.316	-.082	**.836**	.790	.804	.526
9.17 draft stiffness item 17	**.535**	-.057	.061	.382	**.772**	.644	.635	.704
11.35 draft stiffness item 35	.009	**.737**	.184	.092	**.775**	.932	.760	.665
8.5 draft stiffness item 5	-.159	**1.010**	-.223	-.100	.386	**.665**	.316	.334
12.3 traditional severity item G	.045	**.716**	.163	.082	.774	**.919**	.748	.654
12.4 traditional severity item C	.144	**.735**	-.039	.100	.741	**.882**	.645	.639
11.34 draft stiffness item 34	.046	**.604**	.307	.086	.803	**.912**	.815	.662
6.2 traditional duration item D	-.100	**.744**	.149	.002	.591	**.774**	.600	.497
12.1 traditional severity item A	.192	**.617**	.035	.147	.785	**.884**	.697	.678

Items	Rotated component loadings							
	Pattern matrix				Structure matrix			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
11.36 draft stiffness item 36	-.042	**.563**	.419	.087	.771	**.884**	.831	.642
7.2 draft stiffness item 2	.025	**.552**	-.202	.410	.544	**.691**	.422	.673
7.3 draft stiffness item 3	.214	**.513**	-.293	.373	.609	**.709**	.430	.678
11.37 draft stiffness item 37	-.084	**-.434**	-.111	.222	-.370	**-.435**	-.367	-.165
10.28 draft stiffness item 28	-.209	-.007	**.981**	.082	.596	.579	**.857**	.463
10.29 draft stiffness item 29	-.003	-.178	**.916**	.127	.647	.549	**.854**	.489
10.27 draft stiffness item 27	.107	.049	**.722**	.089	.759	.700	**.886**	.563
10.31 draft stiffness item 31	.284	-.087	**.730**	-.057	.747	.611	**.858**	.441
10.32 draft stiffness item 32	.477	-.091	**.557**	-.035	.817	.649	**.844**	.489
10.30 draft stiffness item 30	.422	.120	**.475**	-.044	.857	.754	**.865**	.539
7.4 draft stiffness item 4	-.176	-.217	.154	**.925**	.339	.349	.347	**.759**
8.8 draft stiffness item 8	-.087	.063	.128	**.711**	.494	.543	.477	**.766**
8.7 draft stiffness item 7	.134	.155	.033	**.558**	.618	.639	.537	**.756**
9.9 draft stiffness item 9	.348	.150	.047	.330	.701	.663	.595	**.663**

NB **bold loadings** = highest loading for that item; underlined loadings = other loadings ≥.4

NLPCA

Items	Rotated component loadings							
	Pattern matrix				Structure matrix			
	1	2	3	4	1	2	3	4
9.11 draft stiffness item 11	**1.062**	-.062	-.184	-.007	**.875**	.548	.587	.410
9.10 draft stiffness item 10	**.981**	.028	-.220	.103	**.881**	.603	.592	.504
9.21 draft stiffness item 21	**.711**	.178	.086	-.160	**.824**	.665	.695	.291
9.12 draft stiffness item 12	**.779**	.045	.100	-.027	**.874**	.645	.719	.405
10.25 draft stiffness item 25	**.881**	-.123	.007	-.041	**.780**	.478	.583	.334
9.13 draft stiffness item 13	**.770**	.131	.022	.006	**.882**	.685	.707	.441
9.20 draft stiffness item 20	**.807**	.054	-.003	-.085	**.801**	.578	.622	.326
9.16 draft stiffness item 16	**.780**	.000	-.141	.278	**.806**	.563	.563	.601
10.33 draft stiffness item 33	**.650**	-.136	.359	-.058	**.803**	.539	.742	.333
10.24 draft stiffness item 24	**.599**	.068	.278	-.006	**.857**	.675	.783	.416
10.23 draft stiffness item 23	**.522**	.101	.360	-.132	**.805**	.657	.781	.298
10.26 draft stiffness item 26	**.653**	-.074	.205	.050	**.782**	.542	.674	.410
10.22 draft stiffness item 22	**.609**	.241	.067	-.235	**.716**	.613	.613	.185
9.18 draft stiffness item 18	**.505**	.156	.295	.029	**.855**	.723	.802	.449
9.15 draft stiffness item 15	**.550**	.005	.301	.100	**.833**	.638	.765	.481
9.19 draft stiffness item 19	**.469**	.173	.379	-.089	**.838**	.724	.825	.353
9.17 draft stiffness item 17	**.522**	.141	.027	.268	**.770**	.636	.627	.589
11.35 draft stiffness item 35	-.008	**.790**	.219	.000	.710	**.936**	.758	.411
8.5 draft stiffness item 5	.154	**.791**	-.487	.010	.337	**.568**	.182	.235
12.3 traditional severity item G	.016	**.845**	.143	-.017	.706	**.948**	.732	.400
12.4 traditional severity item C	.032	**.837**	.073	-.007	.667	**.906**	.672	.388
11.34 draft stiffness item 34	.082	**.621**	.317	.017	.765	**.903**	.814	.437
6.2 traditional duration item D	.056	**.658**	.125	.035	.628	**.799**	.636	.387
12.1 traditional severity item A	.043	**.843**	.106	-.018	.703	**.939**	.714	.398

Items	Rotated component loadings							
	Pattern matrix				Structure matrix			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
11.36 draft stiffness item 36	.002	**.494**	.497	.034	.744	**.852**	**.852**	.430
7.2 draft stiffness item 2	-.158	**.961**	-.070	.093	.502	**.842**	.507	.395
7.3 draft stiffness item 3	-.019	**1.053**	-.276	-.042	.483	**.832**	.420	.288
11.37 draft stiffness item 37	.098	**-.529**	-.365	.130	-.488	**-.658**	-.605	-.183
10.28 draft stiffness item 28	-.063	-.127	**.929**	.144	.631	.530	**.847**	.410
10.29 draft stiffness item 29	.090	-.200	**.847**	.120	.660	.498	**.824**	.399
10.27 draft stiffness item 27	.135	.060	**.714**	.068	.759	.676	**.885**	.428
10.31 draft stiffness item 31	.393	-.222	**.685**	.000	.765	.524	**.834**	.354
10.32 draft stiffness item 32	.500	-.115	**.544**	-.033	.822	.595	**.837**	.365
10.30 draft stiffness item 30	.499	.023	**.481**	-.040	.866	.686	**.866**	.392
7.4 draft stiffness item 4	-.114	-.069	.138	**.858**	.358	.308	.326	**.826**
8.8 draft stiffness item 8	.003	.056	.144	**.725**	.503	.463	.459	**.805**
8.7 draft stiffness item 7	.151	.374	.027	**.401**	.625	.666	.552	**.641**
9.9 draft stiffness item 9	**.417**	.200	.043	.250	.710	.625	.596	**.551**

NB **bold loadings** = highest loading for that item; underlined loadings = other loadings ≥.4; *****Specifying a 4 component ordinal solution

Appendix Z: Scoring the RAST

Scoring instructions

The RAST can be scored as individual components or as a total scale. To score, convert each item score to a % score using the tables below. To generate a % score for individual components, add the % scores together and divide by the number of items within the component. To generate a % score for the total scale add together each % score for each individual component and divide by three.

Missing data

One missing item per component is acceptable. In this case, individual component scores can be generated by adding together the item % scores from the available data and dividing that by the number of item % scores provided (one less than the component total). A score for each individual component is required to generate a % score for the total scale.

Severity component			
Item no.	**Score as %**		**Item % score**
1	0=0% 1=25% 2=50%	3=75% 4=100%	
2	0=0% 1=33%	2=67% 3=100%	
3			
4	0=0% 1=10% 2=20% 3=30% 4=40% 5=50%	6=60% 7=70% 8=80% 9=90% 10=100%	
5			
6			
7	0=0% 1=17% 2=33% 3=50%	4=67% 5=83% 6=100%	
8	0=0% 1=10% 2=20% 3=30% 4=40% 5=50%	6=60% 7=70% 8=80% 9=90% 10=100%	
Severity % score (item 1 % score + item 2 % score + item 3 % score + item 4 % score + item 5 % score + item 6 % score + item 7 % score + item 8 % score / 8)			%

Physical component		
Item no.	**Score as %**	**Item % score**
9	0=0% 1=33% 2=67% 3=100%	
10		
11		
12		
13		
14		
15		
16		
Physical % score (item 9 % score + 10 % score + item 11 + item 12 % score + item 13 % score + item 14 % score + item 15 % score + item 16 % score / 8)		**%**

Psychosocial component		
Item no.	**Score as %**	**Item % score**
17	0=0% 1=33% 2=67% 3=100%	
18		
19		
20		
21		
Psychosocial % score (item 17 % score + item 18 % score + item 19 % score + item 20 % score + item 21 % score / 5)		**%**

RAST total % score (Severity % score + Physical % score + Psychosocial % score / 3)
%

Appendix AA: Final layout of the RA stiffness PROM

Please tell us about your RA stiffness

This questionnaire is about RA stiffness that comes and goes. It is not about joints that are permanently stuck (for example, due to an operation). However, we do appreciate that sometimes even permanently stuck joints do get stiffer (for example, when your disease is bad). Please just try think about the stiffness that comes and goes as you answer this questionnaire.

Severity component

This section asks about the severity of your RA stiffness.

1. Over the past 7 days when have you experienced **RA stiffness**? Please tick all that apply to you

   - In the night ☐
   - In the morning ☐
   - In the afternoon ☐
   - In the evening ☐
   - None of these ☐

For each of the following questions, please tick the one answer that best applies to you

2. Have you experienced **RA stiffness** in your joints over the past 7 days?

   - No, not in any of my joints ☐
   - Yes, in a few of my joints ☐
   - Yes, in many of my joints ☐
   - Yes, in all of my joints ☐

3. Over the past 7 days have you experienced **RA stiffness** after a period of immobility (for example, after sitting for a while)?

   - No, not at all ☐
   - Yes, a little ☐
   - Yes, quite a lot ☐
   - Yes, very much ☐

4. Please circle the number that best describes the severity of your **RA stiffness** over the past 7 days

No stiffness 0 1 2 3 4 5 6 7 8 9 10 Extreme stiffness

5. Please circle the number that best describes the impact that **RA stiffness** has had on your life over the past 7 days

No impact at all 0 1 2 3 4 5 6 7 8 9 10 A great deal of impact

6. Please circle the number that best describes how important **RA stiffness** has been in your life over the past 7 days

Not important at all 0 1 2 3 4 5 6 7 8 9 10 Very important

The following questions are often used to assess stiffness. You may be familiar with them from clinic or other questionnaires.

Question 7 asks about your stiffness today.

7. Were your joints stiff when you woke up today? No ☐ Yes ☐
   If yes, how long did this extra stiffness last?
   - ☐ Less than 30 minutes
   - ☐ 30 minutes to an hour
   - ☐ 1 - 2 hours
   - ☐ 2 - 4 hours
   - ☐ More than 4 hours but less than all day
   - ☐ All day

Please answer question 8 in the most appropriate way for you.

8. How would you describe the overall level of morning stiffness you have had from the time you wake up?

No stiffness 0 1 2 3 4 5 6 7 8 9 10 Very severe stiffness

Physical component
This section asks about the physical and daily life impact of your RA stiffness.

	Not at all	A little	Quite a lot	Very much
9. Has **RA stiffness** made it difficult to dress or undress yourself?				
10. Has **RA stiffness** made it difficult to wash yourself (for example, have a shower)?				
11. Has **RA stiffness** made it difficult to carry out your responsibilities or commitments?				
12. Has **RA stiffness** made it difficult to do your daily tasks or activities?				
13. Has **RA stiffness** made it difficult to get out of bed?				
14. Has **RA stiffness** made it difficult to do fine movements (for example, write with a pen)?				
15. Has **RA stiffness** made it difficult to grip or hold things?				
16. Has **RA stiffness** made it difficult to balance without physically supporting yourself?				

Psychosocial component
This section asks about the psychosocial impact of your RA stiffness.

	Not at all	A little	Quite a lot	Very much
17. Have you felt frustrated because of **RA stiffness**?				
18. Have you felt worried or concerned because of **RA stiffness**?				
19. Have you felt self-conscious because of **RA stiffness**?				
20. Have you had to change your plans or behaviour because of **RA stiffness**?				
21. Have you had to work around your **RA stiffness** (or do things in a different way)?				

Thank you for completing this questionnaire

Appendix BB: Abstract (Study 1)

Patients' experience of stiffness in RA are more than just duration and severity
Halls, S., Dures, E., Kirwan, J., Pollock, J., Baker, G., Edmunds, A., Hewlett, S. (2014)

Background: Stiffness is commonly experienced by patients with rheumatoid arthritis (RA). It has considerable impact on their daily lives and influences decisions to seek medication review. Traditionally, stiffness is evaluated by severity and duration, yet research into how patients experience the symptom is limited.

Methods: Patients were purposefully sampled from out-patient rheumatology clinics to reflect a range of age, gender and disease duration. Semi-structured interviews were conducted according to a piloted interview guide. Interviews were analysed using thematic analysis, with a subset analysed by the research team including patient partners.

Results: 16 patients (5 male, 11 female) aged 33 - 78 years (mean 57.3) with disease durations 1 - 27 years (mean 11.5) participated. Analysis identified six themes, each of which fitted around the central concept of 'I experience stiffness as...'

'Part of my disease': Stiffness was considered a normal consequence of RA ("it's really one of the most obvious symptoms of the condition") and was influenced by disease related aspects such as flare and damaged joints ("once you've got damage, you're always stiff"). Relationships to other symptoms such as pain were apparent, but patients could discuss stiffness independently. For some stiffness was a significant symptom in early disease ("stiffness was absolutely integral to the definition of the disease at that time").

'Part of my behaviour and environment': Movement was a key influence, and related to immobility ("you are stiff after being laid in bed") and over-activity ("if I have had like a busy day, and I haven't been able to rest then I might find that it is creeping back in the evening"). Patients also highlighted medications, weather and diet.

'Located within my body': Stiffness was a bodily experience that affected the joints ("there is actual joint stiffness"). Location varied and was reported by some as more widespread during the morning or during a flare ("it affects more joints than it does when I'm not so bad").

'Having consequences': Patients defined stiffness by impact on a range of domains including physical function ("I could not get my fingers to pick up this blooming stupid screw"), quality of life and wellbeing.

'Needing to be managed': Patients managed stiffness using movement, heat and cold, medications, gadgets, and behavioural strategies.

'Variable': Stiffness varied within and between patients and was compounded by the fluctuating nature of RA e.g. stiffness during a flare was "an exaggeration of itself". Additionally, there was temporal variability ("on a good day it is really just morning and evening").

Conclusions: Patients' experiences of stiffness were varied, complex and not fully captured by severity and duration. Future research directions include using these findings to develop a more patient oriented measure of stiffness which might evaluate different dimensions of the symptom.

Keywords: Stiffness, rheumatoid arthritis, patient experience

Appendix CC: Abstract (Systematic literature review)

A systematic literature review (SLR) of the measurement of stiffness in patients with rheumatoid arthritis (RA)

Serena Halls, John Kirwan, Sarah Hewlett

Background: Morning stiffness was omitted from the RA core set because of poor measurement properties of available patient reported outcome measures (PROMs)[1], yet it remains a frequently used clinical and research outcome[2] and is an important symptom to RA patients[3,4,5]. In recent qualitative work, patients highlighted stiffness reduction as crucial to consider themselves in remission[6]. Additionally, an international patient and healthcare professional Delphi, prioritized stiffness (79% consensus)[7] and included it as a core domain for flare assessment (91% consensus)[8]. In both remission and flare, the assessment of stiffness has been identified as an important area of investigation[8,9]. An SLR investigating stiffness PROMs for use in the assessment of RA patients in low disease activity or remission states concluded that there was insufficient evidence regarding stiffness assessment in that context[9].

Objectives: This SLR aimed to expand and update the previous SLR[9] of stiffness in RA remission, to identify the current stiffness PROMs available for RA in general and summarise the evidence of their measurement properties.

Methods: To update the previous SLR[9] an extensive PubMed database search was performed for dates 20/11/12 to 22/09/15, but not limited to remission. Article screening and data extraction were performed using an approach that was consistent with the original SLR[9] by multiple researchers. To expand the previous SLR[9], data extraction was also performed by one author on the 16 articles identified in the original review (14 of which were previously excluded as they did not provide data on stiffness assessment in remission).

Results: In the updated search 147 articles were identified, from which 23 full text articles were screened and 9 included. The 16 articles identified in the original SLR[9] were also included, totalling 25 articles. The fifty identified PROMs predominantly assessed stiffness from the two concepts of severity/intensity or duration, and focused on morning or early morning stiffness alone. Despite covering so few concepts, there was great variation in these PROMs in relation to wording, response options, format and timeframe, and many items were poorly described. Reports of face, content, criterion and construct validity, and stability and sensitivity were minimal.

Conclusions: Current RA stiffness assessment is varied, poorly defined and does not appear to have been developed according to PROM development guidelines[10]. Importantly it is also inconsistent with the patient perspective[4,5] of this symptom. Further work is required to investigate the most appropriate way to assess stiffness in an RA population.

References: [1]Felson et al, 1993; [2]Kalyoncu et al, 2009; [3]Hewlett et al, 2005; [4]Halls et al, 2014; [5]Orbai et al, 2014; [6]van Tuyl et al, 2015; [7]Bartlett et al, 2012; [8]Bykerk et al, 2014; [9]van Tuyl et al, 2014; [10]FDA, 2009

Appendix DD: Abstract (Study 4 development)

Developing a new rheumatoid arthritis (RA) stiffness patient reported outcome measure (PROM)
S. Halls, E. Dures, J. Kirwan, J. Pollock, G. Baker, A. Edmunds, S. Hewlett

Background: Morning stiffness is a frequently used clinical and research outcome measure and is important to patients[1], but was omitted from the RA core set because of poor measurement properties[2]. Current stiffness assessment is inconsistent with patient's perspectives of the symptom[3,4] and has not been developed according to PROM development guidelines[5]. The appropriate content of a new RA stiffness PROM was previously explored and developed through qualitative interview[3] and focus group studies. Draft items were subsequently tested and refined with patients during cognitive interviews, resulting in 39 draft items for inclusion in a PROM. Here we report a quantitative assessment to create the smallest and most internally consistent set of items for a developmentally valid stiffness PROM.

Objectives: To develop the content and structure of a new RA stiffness PROM.

Methods: A postal questionnaire pack was sent to patients with RA based in the South-West of England. It contained 45 items assessing stiffness (39 draft items and 6 items currently used in stiffness assessment), individual items capturing pain (VAS), fatigue (NRS), patient global assessment (VAS), questionnaires capturing disability (MHAQ), and patient-reported disease activity (PDAS26), and basic demographic information. Initial investigation identified items with poor response rates, distributions or correlations for removal. A series of principal component analyses were undertaken with the remaining items, balancing Cronbach's alpha for internal consistency, stability of the component structure (assessed by multiple analyses using random 50% samples of the respondents (bootstrapping)) and parsimony. Based on the statistical results and aided by expert judgement, the smallest number of informative items were retained.

Results: 277 patients (91 male) aged 23-97 years with disease durations 1-45 years participated in the study (42.9% response rate). Seven of the 45 items were removed during initial item investigation. The remaining 38 items demonstrated high Cronbach's alpha (>0.9). During successive rounds of analytical refinement, 17 items were removed.

After round 5, a 3-component structure emerged which remained consistent for a further 13 rounds of testing item removal, demonstrating stability. These components captured 'stiffness severity', 'physical impact' and 'psychosocial impact'. The overall Cronbach's alpha of the final 21 items was 0.95 indicating a homogenous set of items and bootstrapping further demonstrated the stability of the structure.

Conclusions: A new 21 item, 3-component RA stiffness PROM has been developed based on qualitative work[3] with RA patients to enhance content validity. Further testing is now required to assess the validity, reliability and sensitivity to change of the new RA stiffness PROM.

References: [1]Hewlett et al, 2005; [2]Felson et al, 1993; [3]Halls et al, 2014; [4]Orbai et al, 2014; [5]FDA, 2009; [6]Choy et al, 2008; 2015

Appendix EE: Abstract (Study 4 testing)

Construct validity testing of RAST, a new RA stiffness patient reported outcome measure (PROM)

S. Halls, E. Dures, J. Kirwan, J. Pollock, G. Baker, A. Edmunds, S. Hewlett

Background: Morning stiffness is a frequently used clinical and research outcome measure and is important to patients[1], but was omitted from the RA core set because of poor measurement properties[2]. Current stiffness assessment is inconsistent with patient's perspectives of the symptom[3,4] and has not been developed according to PROM development guidelines[5]. A new 21 item, 3-component Rheumatoid Arthritis Stiffness questionnaire (RAST) has been developed based on qualitative work[3] with RA patients and statistical assessment to enhance content validity and now requires testing for appropriate relationships with other measures of disease (construct validity).

Objectives: To perform construct validity testing of RAST.

Methods: The 21 item RAST was developed from 45 items assessing stiffness included in a questionnaire pack which was posted to patients with RA based in the South-West of England. The questionnaire pack also contained individual items capturing pain (VAS), fatigue (NRS), patient global assessment (PtG VAS), questionnaires capturing disability (MHAQ), and patient-reported disease activity (PDAS2)[6], scores generated from an algorithm including PtG, swollen joint count (SJC), and MHAQ), and basic demographic information. The RAST was subjected to construct validity testing using a correlation matrix of Spearman's correlation coefficients (reported as explained variance (R^2)).

Results: 277 patients (91 male) aged 23-97 years with disease durations 1-45 years participated in the study (42.9% response rate). The individual components and sum score of RAST demonstrated appropriate relationships with other measures of disease (pain R^2=45-72%; fatigue R^2=43-55%; PtG R^2=48-68%; MHAQ R^2=55-78%; PDAS26 R^2=56-78%). The shared variance indicated that while RAST (individual components and sum score) is related to other measures of disease, it is not measuring the same thing. Importantly, as identified by patients in earlier qualitative work[3], the variance explained by the RAST (individual components and sum score) and the (patient-reported) SJC included within the PDAS2[6] was between 31-41% indicating that RAST may capture something not currently included within disease activity assessment. The pattern of relationships between individual components and measures of disease also provided support for the 3-component structure. As expected, the 'physical component' shared the most variance with disability (R^2=78%), the 'severity component' shared the most variance with pain (R^2=72%) and PtG (R^2=68%), and the 'psychosocial component' shared less variance with the above disease related measures.

Conclusions: During preliminary validity testing RAST, the new RA stiffness PROM, demonstrated promising construct validity. Further testing of RAST is now required in a fresh population to generate measurement property evidence of reliability and sensitivity to change to support its use.

References: [1]Hewlett et al, 2005; [2]Felson et al, 1993; [3]Halls et al, 2014; [4]Orbai et al, 2014; [5]FDA, 2009; [6]Choy et al, 2008; 2015